

PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA
MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH
MOHAMED KHEIDER UNIVERSITY – BISKRA
FACULTY OF LETTERS AND FOREIGN LANGUAGES
DEPARTMENT OF ENGLISH LANGUAGE AND LITERATURE



**The Use of Integrated Performance Assessment in Enhancing English
Foreign Language Learners' Communicative Competence**

The Case of First Year Students of English at M'sila University

**Thesis submitted to the Department of English Language and Literature in Partial
Fulfilment of the Requirements for the Degree of Doctorate in EFL Teaching /Learning
and Assessment**

Submitted by:

Cheriet Henen

Supervised by:

Prof. Ahmed Chaouki HOADJLI

Board of Examiners:

Prof. Saliha CHELLI	President	Mohamed Khider University - Biskra
Prof. Ahmed Chaouki HOADJLI	Supervisor	Mohamed Khider University - Biskra
Prof. Ramdane MEHIRI	Internal Examiner	Mohamed Khider University - Biskra
Dr. Manel TRIKI	Internal Examiner	Mohamed Khider University - Biskra
Dr. Salim OUNIS	External Examiner	Abbes Laghrour university- Khenchela
Dr. Farid CHAIRA	External Examiner	Larbi Ben M'hidi University-Oum el Bouagh

Academic Year: 2022/2023

Dedication

To my beloved family and close friends ...

Acknowledgements

First and foremost, I would like to express my sincere gratitude and thanks to my supervisor Prof. Hoadjli Ahmed Chaouki for his valuable and appreciated supervision, guidance, assistance, and, above all, his encouragement and support.

I would like also to deeply thank my respectful Examiners: Prof. Saliha CHELLI, Prof. Ramdane MEHIRI, Dr. Manel TRIKI, Dr. Salim OUNIS, and Dr. Farid CHAIRA for their time and effort in evaluating and helping to enhance my work.

Additionally, I would like to profoundly thank all major and minor contributors that helped me in the achievement of my PhD research including the teachers and students who participated in this study for their priceless help.

I would also like to thank my teacher Dr. Cheriet Imane for the help and support that lasted all along my PHD journey.

Abstract

This study aimed, initially, to examine the effectiveness of the Integrated Performance Assessment (IPA) in assessing and promoting students' language proficiency. Additionally, it attempted to investigate current assessment perceptions of and practices among Algerian English as a Foreign Language university teachers (EFL) in an attempt to understand the oral expression teaching and assessment practices being the selected course to conduct the experiment. Furthermore, it sought to examine the washback effect of the IPA on EFL students' motivation and learning. Accordingly, an exploratory sequential Mixed-methods approach was opted. First, an exploratory descriptive method was used through the administration of an online questionnaire in seeking to investigate the current oral expression teaching and assessment perceptions of and practices among Algerian EFL university teachers. Second, and attempting to test the effectiveness of using the IPA in improving language proficiency, a quasi-experimental pre- and post-test control group design was conducted. Finally, another questionnaire was directed to the students of the experimental group to examine the washback effect of this assessment tool on their motivation and learning. This study involved a sample of thirty-five teachers from different Algerian universities, fifty-two EFL first-year students at M'sila University, in addition to twenty-five students for the piloting stage. The findings revealed that Algerian EFL teachers hold positive attitudes towards the use of authentic and performance-based assessment methods. In addition to that, it was found that the IPA can contribute to the promotion of language skills proficiency among EFL students. Further, the findings confirmed the validity and reliability of the IPA as an assessment tool. Besides, the findings reflected the highly positive attitudes of students toward the implementation of the IPA in their oral expression course having a positive impact on their motivation. Finally, this study highlighted the challenges language teachers may face in integrating the IPA, it evoked more research in this area and suggested more gaps to fill.

Keywords: performance-based assessment, Integrated performance assessment (IPA), language proficiency, language skills, the oral expression course.

List of Acronyms

ACTFL: American Council on the Teaching of Foreign Languages

EFL: English as a Foreign Language

EL: English language

ELT: English Language Teaching

FL: Foreign language

IELTS: The International English Language Testing System

IPA: Integrate performance assessment

L2: Second Language

LMD: License, Master, and Doctorate

PBA: performance-based assessment

PPT: PowerPoint Presentation

SPSS: Statistical Package for Social Sciences

List of Tables

Table 1 The IPA's Three Modes of Communication	72
Table 9. Experts Agreement on Students' Questionnaire.....	113
Table 10. IPA based intervention schedule and allotted times.....	115
Table 11. IPA model structure and objectives.....	116
Table 12. Integrated Performance Assessment Comprehension Guide Analysis: Performance Expectations.	118
Table 13. Integrated Performance Assessment (IPA) Interpersonal Mode Rubric	121
Table 14. Integrated Performance Assessment (IPA) Interpersonal Mode Rubric	123
Table 15. Description of the IELTS Listening test.	126
Table 16. Description of the IELTS Reading test.....	127
Table 17. Description of the IELTS writing test.	128
Table 18. Description of the IELTS Speaking test.....	130
Table 19. Study Phases Description	133
Table 20 Kolmogorov-Smirnov and Shapiro-Wilk normality tests (experimental group)	139
Table 21. Kolmogorov-Smirnov and Shapiro-Wilk normality tests (experimental group)	140
Table 23. Intra-class Correlation Coefficient of Test Raters	143
Table 24. Reliability Analysis of the Raters' Scoring for the Scoring Scale	144
Table 25. Concurrent validity of the IPA test.....	145
Table 26. Objectives of Teaching Oral Expression Module.....	146
Table 27. Four Skills Importance Rank in Oral Expression Session.....	147
Table 28. Oral Expression teaching activities	148
Table 29. Students preferred oral expression teaching activities	149
Table 30. Hindrances faced by teachers when teaching oral expression.....	151
Table 31. Assessment Methods Used by Oral Expression Teachers.....	153

Table 32. Teachers' Feedback Type.....	155
Table 33. Methods of Feedback Provided by Oral Expression Teachers.....	156
Table 34. Types of mistakes targeted by teachers' feedback	157
Table 35. Difficulties Faced by Teachers when Assessing Oral Expression	157
Table 36. Independent groups T-test between experimental and control groups pre-test performance.....	162
Table 37. Descriptives of the Experimental group performance in the pre-test.....	163
Table 38. Distribution of experimental groups pre-test proficiency levels	163
Table 39. Experimental group pre-test performance in the four skills	165
Table 40. Independent unequal groups T-test between the experimental and control groups pre-test performance.....	168
Table 41. Descriptives of experimental and control groups post-test performances.....	169
Table 42. Paired sample T-test between Experimental group pre and post-test performance	172
Table 43. Paired sample T-test between Control group pre and post-test performance...	174
Table 44. Weighted mean intervals	177
Table 45. Students' attitudes towards the oral expression module.....	178
Table 46. Students' attitudes towards IPA activities	181
Table 47. Students' performance in the IPA Descriptives	183
Table 48. Concurrent validity of the IPA test (total scores)	185
Table 49. Concurrent validity of the IPA test (the four skills)	186

List of Graphs

Graph 1 Graph of Normality Distribution Q–Q Plot	141
Graph 2 Objectives of teaching oral expression module graph	147
Graph 3. Four skills importance rank in oral expression session Graph	148
Graph 4 Oral Expression teaching activities graph.....	149
Graph 5. Most enjoyed oral expression teaching activities graph	150
Graph 6. Students' participation in the selection of teaching content graph	151
Graph 7 The use of the four skills in the oral expression session graph.....	152
Graph 8 Assessment Methods Used by Oral Expression Teachers graph	154
Graph 9. Teachers' Feedback Provision Frequency graph	154
Graph 10. Methods of Feedback Provided by Oral Expression Teachers graph	156
Graph 11. The Use of the Four Skills in oral expression assessment graph.....	158
Graph 12. Efficiency of the Four Skills Integration graph	159
Graph 13. The Use of Traditional Methods graph.....	160
Graph 14. The use of Alternative Assessment Methods graph.....	160
Graph 15. The experimental group proficiency pre-test levels graph	164
Graph 16. Experimental groups' four skills performance graph	166
Graph 17. Differences between experimental and control groups four skill performance in the pre-test graph.....	166
Graph 18. Pre-test proficiency levels distribution among experimental and control groups graph.....	167
Graph 19. Post-test Proficiency Levels' Distribution among the Experimental and Control Group graph.....	170
Graph 20. The differences in the four skills performance in the post-test graph.....	171
Graph 21. Differences in the experimental group pre- and post-test performance graph.....	172

Graph 22. Differences in the pre-and post-test proficiency levels distribution graph 173

Graph 23. Differences in control group pre- and post-test performance graph 175

Graph 24. Differences in the pre-and post-test proficiency levels distribution among control group graph 175

Graph 25. Students Attitudes towards Oral Expression Session graph 180

Graph 26. Students' Attitudes towards Assessment IPA Activities and Outcomes graph 182

Graph 27. Students' performance in the IPA 184

Graph 28 Differences in the pre-and post-test proficiency levels distribution 195

List of Figures

Figure 1 Integrated Performance Assessment (IPA): Connecting Instructional Practices, Standards, and ACTFL Performance Guidelines,	62
Figure 2 Integrated Performance Assessment Units: A Cyclical Approach	65
Figure 3 Characteristics of a Descriptive and Co-Constructive Approach to IPA Feedback	74
Figure 4 Discourse Features used to Facilitate a Co-Constructive Approach to IPA Feedback	75
Figure 5 Teachers and Learners as Co-Constructors of IPA Feedback	76
Figure 6 Formative assessment cycle.	94
Figure 7 The relationship between dependent and independent variables	102
Figure 8 Experimental procedures	104

Table of Contents

Dedication	2
Acknowledgements	3
Abstract	4
List of Acronyms	5
List of Tables	6
List of Graphs	8
List of Figures	10
Table of Contents	11

GENERAL INTRODUCTION

1. Background of the Study.....	18
2. Statement of the Problem	21
3. The Research Questions	22
4. The Research Hypotheses	23
5. The Research Aims	24
6. The Significance of the Study	24
7. The Research Methodology	25
8. Structure of the Thesis.....	26

CHAPTER ONE: Assessment in the EFL Classroom

Introduction	29
1.1 Assessment, Testing and Evaluation.....	29
1.2 Assessment in Language Learning.....	31
1.3 Assessment in EFL teaching and learning	37
1.4 Types of Assessment.....	38
1.4.1 Diagnostic Assessment	38

1.4.2 Summative Assessment	39
1.4.3 Formative assessment	39
1.4.4 Assessment for Learning	40
1.5 Assessment Feedback.....	42
1.6 The Performance Based Assessment	43
1.6.1 Definitions	45
1.6.2 Authenticity in the PBA.....	47
1.6.3 Tasks of the Performance-Based Assessment.....	48
1.6.4 The Performance-based assessment variation	49
1.7 The Integrated Performance Assessment as an Alternative	50
1.8 Assessment in Higher Education	51
1.8.1 Assessment in the Algerian Higher Education	52
Conclusion.....	54

Chapter Two: The Integrated Performance Assessment and EFL Proficiency

Introduction	57
2.1 Integrated Performance Assessment	57
2.2 The Nature of the Integrated Performance Assessment (IPA).....	61
2.3 Tasks of the IPA.....	66
2.3.1 The Interpretive Mode	66
2.3.2 Interpersonal Task.....	69
2.3.3 The Presentational Mode	70
2.4 The Three Modes of Communication	71
2.5 The co-constructed feedback of the (IPA)	73
2.6 The Implementation of the IPA in Foreign Language classrooms.....	77
2.7 Scoring Rubrics of the IPA	80

2.7.1 IPA Interpretive Rubric	81
2.7.2 IPA Interpersonal Rubrics	82
2.7.3 IPA Presentational Rubric.....	83
2.8 The importance of implementing the IPA in Language Classrooms	83
2.9 Assessing and Enhancing EFL learners' Language Proficiency	86
2.9.1 Language proficiency	87
2.9.2 Proficiency Levels	90
2.9.3 Assessing proficiency	93
2.10 Developing students' language proficiency through assessment.....	93
Conclusion.....	96

Chapter Three: Research Methodology, Data Collection Instruments, and Procedures

Introduction	98
3.1 Research Design and Methods	98
3.1.1 The Experimental Design	100
3.1.1.1 The Independent and Dependent Variables	100
3.1.1.2 The Pre- and Posttest Control Group Design.....	102
3.2 Research Setting and Participants	104
3.3 Data Collection Instruments.....	107
3.3.1 Teachers Perceptions of and Practices Questionnaire	107
3.3.1.1 Teachers' Questionnaire Reliability and Validity.....	109
3.3.2 Students Questionnaire	112
3.3.3 The IPA Model Based Intervention	114
3.3.3.1 The Interpretive Phase	117
3.3.3.2 The interpersonal Phase	119

3.3.3.3 The Presentational Phase	122
3.3.4 The Language Proficiency Test	124
3.3.4.1 The IELTS Proficiency Test.....	124
3.3.4.2 The IELTS in The pre-test and post-test.....	131
3.4 Procedures	132
3.5 Statistical Tools.....	133
Conclusion	136

Chapter Four: Data Analysis and Interpretation

Introduction	138
4.1 Normality distribution and homogeneity of the research sample	139
4.2 IPA proficiency test reliability and validity	142
4.3 Analysis of Teachers Questionnaire.....	145
4.3.1 Teachers Perceptions and Practices of Oral Expression Teaching.....	146
4.3.2 Teachers Perceptions of and Practices in the Oral Expression Assessment.....	153
4.4 Analysis of the Experimentation Data	161
4.4.1 Data analysis of the first hypothesis	161
4.4.1.1 The description of experimental and control groups proficiency levels in the pre-test	162
4.4.2 Data analysis of the second hypothesis.....	167
4.4.2.1 The description of experimental and control groups proficiency levels in the pre-test.....	169
4.4.3 Data Analysis of the Third Hypothesis.....	171
4.4.4 Data Analysis of the Fourth Hypothesis	173
4.5 Students' Questionnaire Data Analysis.....	176
4.5.1 Students Attitudes towards Oral Expression New Sessions	177

4.5.2 Students' Attitudes towards the IPA Activities and Outcomes	180
4.6 Analysis of students IPA performance.....	183
4.7 Validity of the IPA	184
4.8 Synthesis of the Findings	177
Conclusion	180

**CHAPTER FIVE: Discussion of The Findings, Pedagogical Implications, Limitations,
And Further Research Recommendations**

Introduction	183
5.1 Teachers' Perception of the Integration of the four skills in Language Assessment	183
5.1.1 Teachers' Perceptions of the Oral Expression Module Teaching.....	183
5.1.2 Teachers' Perceptions of the Oral Expression Activities	186
5.1.3 Challenges to achieve proficiency in Language skills	189
5.1.4 The Current Assessment Practices among Algerian EFL Teachers in the Oral Expression Course.....	190
5.2 The Impact of IPA on improving students' language proficiency	192
5.3 The Use of IPA to assess language proficiency	199
5.4 The Washback Effect of the IPA on Students' Attitudes.....	201
5.4.1 Students' attitudes towards the oral expression session after the integration of the IPA	201
5.4.2 Students' Attitudes toward the Integration of the IPA as an Assessment tool	204
5.5 Pedagogical Implications	207
5.6 Limitations and Challenges of the IPA	211
5.7 Suggestions for Future Research.....	215
Conclusion	217

General Conclusion

References	224
Appendices	241
Résumé	280
ملخص البحث	281

GENERAL INTRODUCTION

1. Background of the Study

Assessing the progress of learning is, and will always be, the main concern of educators all around the world because it goes hand in hand with both teaching and learning processes. Since teaching cannot be efficient if not strictly connected to effective assessment, imperfectly selected assessment tasks and tools can have a negative impact on both learners' learning and development (Espinosa, 2015). Therefore, it is a teacher's duty to look for innovative assessment methods that connect assessment with teaching practices so as to improve the learners' language knowledge and skills.

It has been a common practice by teachers, for many years, to rely on standardised tests and traditional ways of assessment to evaluate students' learning. However, related research confirmed that these types of assessment cannot cover all aspects of learners' performance in the target language. This is due to the fact that traditional assessment methods are often seen as indirect, inauthentic, and standardised (Dikli, 2003). In the same vein, more researchers like Law and Eckes (1995) emphasise the same issue confirming that traditional assessments are single-occasion tests that measure what learners can do at a particular time. Therefore, test scores cannot really be a real reflection of a learner's progress, and cannot tell what particular difficulties the learners had faced during test completion (Dikli, 2003).

Moreover, traditional assessments like pencil-and-paper tests are usually reductionist, and do not measure learners' acquisition of the contents, their deeper understanding of a topic, or their mastery of complex skills like critical thinking or problem-solving (Espinosa, 2015). Additionally, they do not provide the necessary input required to reconsider teaching practices in order to improve instructional practices or assist weaker learners in their process of learning (Espinosa, 2015).

Thus, it is becoming unavoidable to put into practice new alternative assessment strategies that consider learners' cognitive development, put emphasis on what learners can do instead of what they cannot do, serve in diagnosing learning and teaching problems, and whose core concern is the continuous improvement of learning for all learners (Bejarano & Gordon, 2009). This has led to the shift from traditional assessment towards alternative forms of assessments through which researchers and educators are trying to find effective assessment tools to evaluate learners appropriately, and that go beyond paper-and-pencil formats (Dikli, 2003). This reforming shift in assessment has resulted in the inception of Performance-Based Assessment (PBA) and Authentic Assessment (AA) as effective forms for assessing students' progress in meeting the standards of language use since they require goal-directed use of language, use of multiple skills or modes of communication, and integration of language skills (Adair-Hauck et al., 2006).

The practice of PBA has recently attracted considerable interest in education literature. In the 1990s, Performance-based Assessments became a legitimate alternative to traditional Multiple-choice Tests (Hammond & Adamson, 2013). According to Reeves (2000), the emphasis on PBA is due to the importance of consideration of learners' ability to apply his/her knowledge and skills to real-life simulations. According to Espinosa (2006), Performance-based Assessments represent a set of strategies that asks students to use their knowledge and skills to create a product or perform a task that is authentic and meaningful to them based on certain predetermined criteria.

Integrated Performance Assessment (IPA), as a variation of PBA practices, is defined as a classroom-based performance assessment model. It is an assessment model that provides a real world context for assessing real world competence in the five goal areas, which are much broader than just linguistic proficiency, by asking the learner to complete an authentic task (Adair-Hauck et al., 2014). It is an assessment strategy that would demonstrate what foreign

language learners could do in the target language (Schenck, 2019). It has gained considerable acknowledgement since its first proposition by the *American Council on the Teaching of Foreign Languages* (ACTFL) that received a U.S. Department of Education International Research and Studies grant in 1997 to design an assessment prototype in response to the prevailing detachment between assessment research and practice, as well as to address the need for a way to assess learner progress in attaining the standards (Glisan et al., 2006).

According to the same source, this model can be used for evaluating learner's communication skills, as well as their proficiency in the three modes of communication (Interpretive, Presentational, and Interpersonal) identified in the World-Readiness Standards for Learning Languages (2015). The World-Readiness Standards incorporate the Standards for Foreign Language Learning: Preparing for the 21st Century (1996), and Standards for Foreign Language Learning in the 21st Century (2006).

The IPA model provides opportunities for learners to demonstrate their ability to communicate using their language skills (Listening, Speaking, Reading, and Writing) around a specific theme across its three modes of communication, as well as the opportunity to receive co-constructive feedback that helps them to improve. Core features of the IPA include its emphasis on a "Cyclical Approach" to assessment, which combines continuous modeling, practice, performance, and feedback, its continuous feedback provision to learners throughout the assessment process, and the use of authentic texts for the interpretive task (Davin et al., 2011). Moreover, IPA implementation studies highlighted the benefits of using this form of assessment since it can: serve as a consciousness-raising technique to influence or encourage teachers to modify or refocus their instructional plans to better meet the needs of the learners, and have a positive "Washback Effect" on learning, as well instruction (Glisan et. al, 2003)

2. Statement of the Problem

Paper-pencil tests and oral presentations have always been the main methods of assessing EFL learners' knowledge and skills in the Algerian universities, simply because the main focus of those tests is evaluating to what extent the teacher achieves the course objectives and covers the courses' contents. However, these tests are widely viewed by researchers as being not adequate to depict the learners' language proficiency. Although current research suggests new paradigms for assessments, virtually no alternative assessment in the Algerian university, as an EFL context, has focused on establishing links between classroom experiences and performance since the assessment's main focus was on the accuracy, grammar, fluency, and modules' contents (Sarnou, et al., 2012).

This dilemma has restricted EFL learners' use of language and their practice opportunities especially during the Oral Expression classes. In this regard, Thornbury (2005) states that learners' failure, in developing communicative skills, is highly associated to a lack of practice opportunities developing instead only learners' vocabulary and grammar. This deficiency results negative affective problems including the lack of motivation and anxiety (p.28). This may lead to question what EFL teachers are really assessing, is it solely language itself or language in use? And how do they help their students to improve their performance in the target language?

To solve such a problem, an alternative approach to assessment called the IPA has been created and designed to assess and develop language learners' proficiency. This new approach to language assessment and teaching has been implemented in foreign language classrooms in the USA and the U.K, as well as some other European countries for several decades. Significant results have been detected after its use, the reason why many language schools and institutions around the world implemented it as both a pedagogical and assessment tool in

order to assess and develop students' abilities and communicative skills to use the language in the fullest meaning of the term (Adair-Hauck, et al, 2006). Therefore, it can be said that such an assessment model is not just an evaluation tool but an educative one. According to Wiggins (1998), tests can be educative in that they could enhance the learners' performance as well as teachers' instruction.

In light of all what has been mentioned, concerning the PBA, in general, and the IPA, in particular, it is possible to anticipate that the IPA model can be one of the most appropriate and effective evaluation strategies that can be used with EFL learners in the Algerian EFL context in order to promote language proficiency. Accordingly, this study is set on the assumption that the use of IPA can have a positive impact to promote EFL teaching and learning equally. Thus, this study is an attempt to examine the effectiveness of using the IPA model to assess learners' proficiency, and to track the outcomes of this use on enhancing EFL students' language skills and motivation.

3. The Research Questions

In accordance with the stated problem, this study holds a number of overarching research questions:

Q1: What are the current assessment practices among the Algerian EFL teachers to assess their learners' language learning in the oral expression module?

Q2: To what extent is the use of IPA model effective in assessing and improving EFL learners' proficiency?

Q3: What is the washback effect of the suggested IPA on students' attitudes towards teaching and assessment in the oral expression module?

In addition to these main questions, a number of sub-questions were also asked in relation to the experimental design. These questions are:

Qa: Is there a statistically significant difference between the experimental and control groups' performance in the proficiency pre- test?

Qb: Is there a statistically significant difference between the experimental and control groups' performance in the proficiency post- test?

Qc: Is there a statistically significant difference between the experimental group performances in the pre- and post-tests?

Qd: Is there a statistically significant difference between the control group performances in the pre- and post-tests?

The fifth question is meant to test the concurrent validity of the IPA

Qe: Is there a statistically significant correlation between the experimental group results in the proficiency test and the IPA model?

4. The Research Hypotheses

In relation with the research questions a number of hypotheses were assumed to guide this study's investigation and data analysis procedures. The main hypotheses of this study, which sought to answer the second question, are stated as follows:

- The use of the IPA model can have an effective impact on enhancing EFL students' language proficiency.
- The use of the IPA can have a positive washback effect on fostering EFL students' motivation and learning.

Along with these research hypotheses, the statistical sub-hypotheses were assumed for the inferential statistical analysis:

H0a: There is no statistically significant difference between the control and experimental groups' performance in the proficiency pre- test.

H0b: There is no statistically significant difference between the experimental and control groups' performance in the proficiency post-test.

H0c: There is no statistically significant difference between the experimental group performances in the pre and post-test.

H0d: There is no statistically significant difference between control group performances in the pre- and post-test.

H0e: There is no statistically significant correlation between IPA score and the proficiency post-test scores.

5. The Research Aims

This study aimed to, first, to explore the current assessment perceptions of and practices among the Algerian EFL university teachers. It also sought to gain an understanding of the impact of these practices on both teaching and learning. Specifically, it attempted to understand oral expression teaching and assessment practices being the selected course to conduct the experiment.

Additionally, the present study aimed to evaluate the effectiveness of using the IPA model in assessing EFL Algerian students' language proficiency, and its impact on both teaching and learning in the oral expression classroom. Furthermore, it attempted to examine the effectiveness of using this alternative assessment model in oral expression sessions to promote the Algerian EFL students' language proficiency.

6. The Significance of the Study

This study could be important as it addresses a gap in language assessment in the Algerian university context that we believe it was overlooked in recent research. Moreover, reviewing the related literature to the use of PBA, in general, and IPA model, specifically,

reveals that this study can be seen as one of the initiatives, in the field of assessment, in the Algerian context since the IPA is an innovative assessment tool that has never been integrated in an Algerian context.

It also gains importance since it casts light on the current language assessment practices among Algerian EFL teachers, which can give insights on the weaknesses, as well as, the strengths of these practices and calls for remedial interventions and further studies to improve them. Moreover, it ends up by designing an IPA model to assess EFL students' language proficiency that can help to enhance their language skills and foster their motivation.

7. The Research Methodology

In the light of the objectives of this study, a Mixed-methods approach for data collection seems to be the most suitable method. Thus, and given the steps undertaken in this study, an exploratory sequential Mixed-methods design was adopted. This design started with a qualitative data collection and analysis stage, followed by a quantitative data collection and analysis stage.

First, in seeking to investigate the current oral expression teaching and assessment perceptions of and practices among the Algerian EFL university teachers an exploratory descriptive method was used through the administration of a questionnaire for data collection among EFL university teachers in various Algerian universities.

Second, and attempting to test the effectiveness of using the IPA assessment model in improving language proficiency, an experimental design was opted. The implemented design in this study was the pre and post-test control group design. The obtained results from both the pre-test and post-test were compared to measure the effectiveness of integrating the IPA model on enhancing the EFL students' proficiency.

Finally, to answer the final research question, another questionnaire was directed to EFL students, who participated in the research experiment, with the purpose of exploring their attitudes towards the use of the IPA in the oral expression session to help them to improve their language proficiency.

The population of this study was the Algerian EFL university students. The participants in this study were conveniently selected from the accessible population of first year students at M'sila University during the academic year 2021-2022.

The study involved 52 students that were divided in two groups, 26 students in the experimental group, and 26 in the control group. In addition, 35 teachers, participated in the qualitative study, exploring their perceptions of and practices in the oral expression course.

8. Structure of the Thesis

The present thesis is divided into 5 chapters, in addition to a General Introduction and a General Conclusion. The chapters are divided as follow:

Chapter one presents a selected review of the related literature on the central research concept of this study which is Assessment, in general, and the performance-based Assessment, specifically. Also, it sheds light on the importance of feedback in formative assessment. Additionally, it provides a brief description of the assessment practices in departments of English in the Algerian universities, as well as the need for an alternative assessment. Then it ends up presenting the IPA as an alternative to traditional assessment practices.

Chapter Two provides a brief description and a review of literature on the study variables "the integrated performance assessment" and the "EFL Students' Language Proficiency". First, it presents the components of this new testing system, then it ends with

a description of the scoring procedures for this model based on the manual.

Chapter Three offers a framework of the research methodology adopted in the practical part of the current study. It describes the research methodology and procedures, research design and data collection methods.

Chapter Four discusses and analyses the results and data obtained from the present study. The chapter ends with a discussion and a summary of the findings.

Chapter Five this chapter presents the discussion of the overall findings of this study. The findings are presented and discussed within the framework of the related literature. Additionally, this chapter sheds light on the study outcomes' and pedagogical implications, then presents all the limitations encountered by the researcher. Finally, this chapter suggests a number of recommendations for further research.

CHAPTER ONE:

Assessment in the EFL Classroom

Introduction

This chapter starts with a general review of assessment that aims to give definitions of the term and provides a brief history of assessment including its importance in the FL Classroom. After that, it steps forward to discuss the current types of assessment in higher education at the departments of English in the Algerian Universities. Then, this chapter goes narrowly to cover Alternative Assessments in FL teaching and learning. Then, the researcher sheds light on the performance-based assessment in the current assessment practices with reference to its importance, authenticity, and various tasks. Finally, this chapter ends with limiting the scope to one type of the PBA model, that is Integrated Performance assessment (IPA), in the EFL classrooms as an alternative.

However, an important distinction should be made, first, between the terms “Assessment”, “Evaluation” and “Testing” so as to avoid any ambiguity or misconception while tackling the core idea related to assessment in our present study.

1.1 Assessment, Testing and Evaluation

The terms: assessment, testing and evaluation are sometimes used interchangeably by researchers and teachers to refer to the same phenomenon in language classrooms that is to measure and collect data about the learning/teaching process in general and the learners' progress specifically, including how well they perform a task or skill, and how well they achieve a specific learning objective (Shankar, 2016). However, these measurement tools are quite different in use within the EFL classrooms. In fact, each one of them has a different use and purpose and can be defined differently.

Brown (2003) distinguished between the terms Assessment, and Tests. He stated that tests are included within assessment and considered as one of its useful tools and procedures.

He defined tests as “Prepared administrative procedures that occur at identifiable times in a curriculum when learners master all their faculties to offer peak performance, knowing that their responses are being measured and evaluated” (Brown, 2003, p. 14).

In other words, tests are procedure applied to measure then evaluate the learner's abilities, performance and skills in a specific period of time and in a given domain while the learners are totally aware that they will be evaluated according to their answers. Whereas, he defined assessment as “An on-going process that encompasses a much wider domain” (Brown, 2007, p445) and it appears through various procedures that can be either oral or written tasks aiming at assessing student's performance all along with the teaching/learning process.

A clear set of distinctions were made between assessment and evaluation by educators and researchers in EFL teaching and learning. According to Rea and Jean (2010), the difference between these two processes is clearly identified through the different use and purpose of each of them.

Unlike evaluation, assessment is a formative ongoing process of “collecting, reviewing and using data” (Yambi & Caluyua, 2020) used to identify both areas of improvement and flaws in the learning process. It includes providing feedback on current performance. Additionally, the main aim of Assessment is to measure learners' knowledge, abilities, and skills in order to make future enhancements to the learning process. Whereas, evaluation is a judgmental summative process “of measuring or observing someone, with an aim of drawing conclusions” (Brown, 2007, p. 446) on the basis of specific criteria and evidence. It tends to provide grades, pass judgments, and ascertain whether the intended objectives are established or not (Rea & Jean, 2010). Therefore, assessment and evaluation differ in that assessment

concerns generally about the process of learning while evaluation's main focus is on the learners' production (Surbhi, 2017).

In this regard, it can be concluded that assessment, evaluation, and testing are three distinct educational key words in language teaching and learning that differ in use and purpose; however, each one of them plays a specific extremely important role in the teaching/learning process.

1.2 Assessment in Language Learning

Assessment has so long been applied as inevitable measurement process in language classrooms. In its early beginnings, assessment was conducted only for summative purposes. In contrast, the growing interest with the formative assessment is a novice concern. Mary James, in her book "An overview of Educational Assessment" asserted that "Assessment for grading, selection and certification (summative purposes) has a two thousand years of history, stretching back to the Han dynasty in China" (James, 2010). The Chinese test, which was the first written examination in history, was conducted once every three years and called as: The Chinese civil service examinations. It was generally based on memorisation, question-and-answer tests, as well as owning a personal notebook for each learner which was called Book of merit (Wilbrink, 1997). This Chinese model of testing was the first source of inspiration for the western world to develop the testing systems, and the first basis on which the present examinations stand.

Throughout the medieval period, education was related to teaching Latin literary and religious sacred texts. The main dominant characteristics of teaching and learning, meanwhile, was teaching grammar and translation through learning by heart and memorisation. Therefore, learners during this era were assessed though recitation and

memorisation of long texts, grammar questions in forms of questions-and-answers, or question each other (*Wilbrink, 1997*).

During the 14th century, between 1370's to 1450's, new innovation in assessment and testing was created and applied by Joan Cele, a famous teacher *Zwolle* (northeastern Netherlands). In the school, where he was working, he came up with a solution for the problems they faced in organising students, and for curriculum division by creating a new strict system of examination. The latter was conducted twice a year by which he selected students who would pass to the next level and classified them into eight different classes. Cele's exam was divided into two parts; the first part included recitation and memorisation, and the second one was all about learners' insights and comprehension of Latin texts. Late in the 14th century, Cele's examination evolved and spread to become a model of examination for European schools and universities. It lasted for so long as a grading and classification tool for learners on the basis of merit. Accordingly, we may conclude that this type of test resembles what is known today as a placement test (*Wilbrink, 1997*).

Early in the 1900's, a new form of test appeared in Europe. It was first presented by the psychologist Alfred Binet and his colleague Theodore Simon to the French government. The test was designed for the French schools to identify learners' weaknesses and classify them according to their needs to provide them with specialised treatments. This new form of school test was not based on what has been taught at school, but on psychological and affective factors that affected the learning process, such as: problem-solving skills, attention, and memory through using questions. Based on Binet's test and its observation, intelligence tests are created today known as The Stanford-Binet Intelligence Test (*Brink, 2011*). Another innovative invention, which was presented early in this period, was The Multiple-choice item. It was created by Frederick J. Kelly in 1914, and was widely adopted during that period. After

research on the new item reflected that its results were more valid, the essay examinations started to be neglected and then abandoned for a long time (Clarke et al., 2010).

By the end of the First World War, examinations and tests remained but took on another dimension. They became widely used in different situations and for various purposes. Among the most popular uses of tests during this period, were the placement Military examinations to choose the right men for leadership and placing recruits in appropriate positions, which were known as the Army Alpha and Beta Tests. Also, IQ tests in universities, for immigrants, for scholarships and promotions of honour (Clarke et al., 2010; *Wilbrink, 1997*).

Between the 1960's and the 1970's, assessment and testing faced a new phase of progression on which new assessment tools and techniques were developed and presented. In this respect, Alan Davies (1982) indicated that: "The Modern Language Testing appeared...Its aim was to introduce teachers to a diversity of testing techniques based on the teaching and testing theories of the mid 1960s" (p.125).

In language learning, this development in assessment practices occurred with the appearance of the structural approach to language teaching and learning. It aimed at using test and examinations to assess learners' proficiency, as well as their communicative skills and abilities.

Thus, the communicative Language testing has attracted the attention of linguists and educators and knew a considerable progress for the value it holds for language learning (Davies, 1982). However, this new testing system was only Competence-based and not taking Performance into account (Skenhan, 1988). Not just that, but the ranking procedures were for so long the only way of giving judgments. Accordingly, educators and instructors provided judgments and comments on the examination papers then ranked the learners from one for the worst (1 to the worst) to the best (number of the learners).

Yet, this ranking system was not sufficient with the increasing demands for objective assessment practices. Thus, Europe has witnessed a shift from this complex grading processes into a new marking system in the French and the English institutions, such as Oxford and Cambridge. At its very beginnings, this ranking system was characterised with the lack of rules and standards, and it was different in each country (*Wilbrink, 1997*). However, with the increasing demands for testing and examinations during this period, new scaling systems appeared in France, England then the USA including the large-scale scoring system (Clarke et al., 2010).

Additionally, the late 1980's, had witnessed an increased interest among testing specialists and educators on tests' efficiency and efficacy. This led them to take an upper step forward in working on test validity, reliability, ethics, politics, computer-related matters and the washback effect (*Wilbrink, 1997*).

Few years Later, from the 1980's to the beginning of the 1990's, remarkable improvements and changes on the field of language assessment and testing have been established in parallel with the development of other areas of Applied Linguistics. As it has been noted by Davies, A. (2013), in his book *Fifty Years of Language Assessment*, stating that the main progress in language testing during the 1980's appeared through the expansion and spread of EFL Tests in several panels and agencies for different purposes, including:

- Communicative Use of English Language (test),
- The Cambridge Examinations,
- Test of English for International Communication (the ETS)
- Test of English as a Foreign Language
- Test of English for Educational Purposes

- British Council's English Language Testing Services (Davies, 2013, p.122)

Moreover, this remarkable progress appeared in many other forms as the spread, and “the increase in books on language testing, both introductory and advanced, as well as the launch of the specialist international journal *Language Testing*” (Skehan, 1989, p. 5).

In the 1930s and 1940s further, Ralph Tyler came to the scene with his new behavioural paradigm. He included a behavioural component in both educational objectives and language testing which paved the way to the appearance of Blooms' Taxonomy of Educational Objectives, and the development of many other works on the field of behavioural objectives of testing (Clarke et al., 2010). According to Taylor's rationale, the educators have to deal with their learners as active participants in the classroom where they take part in all tasks so as to achieve the learning objectives and control the learning experiences through: affording various chances to the learners in order to perform the target behaviours that the curriculum aims to achieve, helping learners to develop and present sufficient behaviour, providing variety and feasibility in both the teaching and learning methods applied in the classroom, offering the learners models of behaviour to help them to identify objectives and use them as models to be imitated, and finally, providing a continuous feedback to the learner as well as the educators (Wraga, 2017).

By the late 1990's, the US educational system developed and implemented the standards-based assessments for more standardised tests, which were followed by the emergence of the alternative assessment movement. The latter gave a rise to new assessment forms that replaced the Multiple-choice questions including the Open-ended questions, Essays, Portfolios and Performance tasks. These new complex forms of assessment were designed to provide more detailed data about the learners' knowledge,

higher order thinking, performance, skills, and the teaching/learning process as a whole (Clarke et al., 2010).

Along with the continuous progress in computer technology in the recent decades, assessment has moved to the virtual world to be web-based and computer-delivered. These new types of assessment were similar in design to paper-pen tests, aiming to measure the same performance, knowledge, and skills and to provide the same results. As Bennett (1999) noted:

Like many innovations in their early stages, today's computerized tests automate an existing process without conceptualizing it to realize the dramatic improvements that the innovation could allow. Thus, these tests are substantively the same as those administered on paper; they measure the same skills, use the same behavioural designs, and depend primarily on the same types of tasks (p.3)

Recently, the growing concern of developing the communicative competence, of the EFL learners, has paved the way to the arrival of a new communicative approach towards EFL teaching and learning. Thus, with the implementation of this approach, there was a need to create a more authentic EFL assessment that would suit the new instruction practices, and in the meantime, effectively measure the EFL learners' skills progress, performance, and knowledge. This need for an alternative assessment appeared as a reaction to the dissatisfaction of the traditional methods of testing. In this regard, *Herman et al. (1992) stated that:*

...growing dissatisfaction with traditional, multiple-choice forms of testing. The result is an explosion of interest in alternative forms of assessment combined with attempts across the country at all levels national, state, local, and classroom to create

them. Talk of portfolios, exhibits, hands-on experiments, and writing-across-the-curriculum abounds (p.254).

Thus, language learning has witnessed a shift towards new forms of assessment which are based on learners' use of knowledge, language abilities, and skills in real world context and authentic situation (McKay, 2006). This alternative is the Performance based assessment that shifted the focus from assessing only learners' knowledge, as the language main concepts and rules, into balancing between the importance of learners' knowledge, and performance that has been neglected for so long. In other words, while the traditional standardised assessments are often conducted to test what learners know by answering questions and quizzes correctly, the PBA has come to test and assess how they use the acquired language and aims at developing "a practical rather than a theoretical command of the language" (McNamara, 2001, p. 334).

In the assessment literature, Alternative Assessment is also known as the Performance Assessment, Performance-based Assessment, as well as the Authentic Assessment. In spite that the PBA, as an alternative, could provide valuable and deep data about the learners' actual levels of language proficiency, skills, and performance; still, the traditional assessments, including the multiple-choice tests, serve as an effective tool to determine the extent to which learners have acquired the basics of the target language. Therefore, assessment specialists and educators found that a mixed approach to assessment could be more efficient in EFL classrooms (*Herman et al.*, 1992).

1.3 Assessment in EFL teaching and learning

Assessment is unquestionably one of the most important practices in the EFL classroom. It plays an integral role in the Teaching/ learning process. Therefore, as language teachers, we have to make sure that the quality of assessment methods we use in our

classrooms are effective enough to best support and help to enhance our learners' learning process.

In the educational context, the term “**assessment**” is an ongoing process of collecting data and information about the learners' performance, skills, knowledge, and needs throughout the use of a set of classroom procedures, and tools in order to measure and evaluate the quality of both teaching and learning progress. In that sense, Richards and Schmidt (2002, p. 35) defined assessment as: “*a systematic approach to collecting information and making inferences about the ability of a student or the quality or success of a teaching course on the basis of various sources of evidence. Assessment may be done by test, interview, questionnaire, or observation*”.

1.4 Types of Assessment

Assessment practices in EFL classrooms include three types. Each type is used for specific reasons at a specific period of time in the teaching/learning process. These three forms of assessment are presented in what is coming:

1.4.1 Diagnostic Assessment

Diagnostic assessment is an evaluation process that takes place before teaching, at the beginning of a course, a year or a semester to diagnose learners' current knowledge, strengths, and weaknesses (William & Thompson, 2007). It helps teachers to obtain many other valuable insights about their learners. These insights can help teachers to:

- bend activities' level of difficulty to what suits learners' levels,
- plan for remedial sessions and additional support to help learners to overwhelm their weaknesses,
- adapt the content of the curriculum and the courses,

- assist learners to overcome their weaknesses and reinforce their strengths through feedback (Ciofalo & Wylie, 2006; Black & Wiliam, 2009; Gurel et al., 2015).

1.4.2 Summative Assessment

The summative assessment, which is an assessment of learning, is a product-based evaluations tool that measures what learners have learned and what they have grasp with the aim of passing judgments (Brown, 2014). Therefore, it is fairly judgmental and not conducted for any improvement since it is grade-based and not feedback-based assessment practice (Cizek, 2010, p. 5). It usually appears at the end of a course, a unit of instruction, a programs or a semester to evaluate students' academic achievements (Brown, 2014).

In the Algerian EFL classrooms, summative assessment is usually restricted to traditional standardised paper-pen examination that includes multiple choice question (MCQ), gap filling, texts comprehension followed by questions and essay writing and is usually conducted under controlled conditions.

1.4.3 Formative assessment

Formative assessment, or the so called assessment for learning, has been defined by many researchers and educators differently. However, all researchers convene in defining it as a “process” rather than a “product” because this type of assessment goes along with the learning/teaching process and it involves different steps.

According to Brown (2004) and Cowie and Bell (1999), formative assessment is an ongoing process of collecting evidence and data about the learners so as to evaluate their actual competencies, skills aiming at enhancing the language learning process and ensure a continuous progress though providing regular feedback during learning.

Additionally, formative assessment helps to compare the learning with the intended goals then all the learning/teaching process is adjusted accordingly (Greenstein, 2010). Hence, it is regarded as a goal-oriented process as it helps to achieve the intended goals during the learning/teaching process.

In the EFL classrooms, the use of various types of assessment would reveal more the actual level of the learners, and provide more information about the learning process. Among classroom tasks, such as: course assignments, observations, reports, project works, presentations, oral questioning, class discussions, quizzes, tests, and daily evaluations (Ruiz-Primo, et al., 2007). Thus, formative assessment entails Performance-based Based Assessment tasks, including all classroom authentic activities and tasks that involve the learners in active learning.

1.4.4 Assessment for Learning

Assessment for learning, or formative assessment, has gained more interest of educators and researchers in EFL learning than other types for its significant role in developing learners' skills and performance (Masuku et al., 2020). In higher education, specifically, assessment for learning has become a widely spread and internationally recommended approach to assessment by various universities for the benefits it brings to learners (Knight, 1995).

Since the main aim in teaching EFL in higher education is to develop learners' language proficiency by providing learning opportunities in order to put into practice the acquired knowledge (Dall' Alba & Barnacle, 2007), therefore, assessment for learning is considered as the most effective pedagogical and evaluative approach to assessment because it provides learners with ample opportunities to learn, practise, find flaws in their performance then ameliorate it through the use of effective feedback (Masuku et al., 2020).

According to Black et al. (2004), assessment for learning is: “any assessment for which the first priority in its design and practice is to serve the purpose of promoting students’ learning” (p. 10).

A number of researchers (Gibbs 2006; Sambell et al., 2013; Boud et al., 2010; William, 2011) identified various characteristics and features to assessment for learning summarised below:

- Authentic contexts and appropriate tasks that grasps all students’ efforts, interests and time;
- Well-designed lesson plan that allocates students’ efforts all over the course timing.
- providing ample opportunities to learners to practice.
- A proper use of summative and formative assessment;
- Offering opportunities for self-evaluation;
- Providing effective feedback including both formal and informal feedback as well as peer feedback to assist student enhance future performance.
- involving all students in productive tasks.

The implementation of this assessment approach in EFL classrooms requires several strategies that can ensure its effectiveness in enhancing learners’ learning. According to Thompson (2008), there are four strategies to follow for an effective implementation of assessment for learning. These strategies are summarised as follows:

- **Create productive assessment tasks:** Types of tasks should be developed in an authentic context which helps to gain outcomes relevant to future professional life.

- **Provide Effective feedback:** It could help student to develop self-regulated learning. Teachers in this type of assessment helps students to figure out their own flaws, create solutions and enhance their performance using feedback.
- **Ensure student Comprehension:** Helping students to understand the requirement of the tasks as well as the expected criteria of their performance in rubrics contributes in preparing students for self and peer-evaluation.
- **Provide opportunities to students to make judgments:** Making judgement includes judgements for both peer and one's own work. It Can be more advantageous and effective than receiving teachers' feedback since it provokes students cognitive and affective processes including: motivation, higher-order thinking, and problem solving.

1.5 Assessment Feedback

Feedback refers to comments, remarks or any information provided to someone who performs about his performance in order to enhance it. As Wang (2006) defines it, "feedback is the information about current performance that can be used to improve future performance" (p.42).

In EFL classrooms, feedback is a direct link that transmits information right from the instructor into the learners about their performance (Boud & Molloy, 2013). It is an integral part of formative assessment on which teachers spot light on key aspects of learners' performance to be improved or reinforced (Klimova, 2015). In the same line of thought, this researcher explained formative assessment as a type of "assessment that is specially intended to generate feedback on performance to improve and accelerate learning" (Nicol & Macfarlane, 2006, p.4). In other words, the formative assessment as an evaluation process is conducted not only to measure and evaluate, but also to provide constructive feedback that

would help to enhance the learners' learning process and achieve the intended objectives (Popham, 2008).

In fact, feedback plays a crucial role in improving learners' skills and performance. It helps to enhance the learning process as a whole (Sambell et al., 2013). In this vein, Brown (2004) identified that teachers provide feedback so as to develop the learner's language abilities by offering learners remarks or propositions, or guiding their attentions to committed errors. Accordingly, effective feedback is the one that "helps students trouble-shoot their own performance and self-correct; that it helps them to take action to reduce the discrepancy between their intentions and the resulting effects" (Nicol, & Macfarlane, 2006, p.5).

Effective feedback has many characteristics. According to Hattie and Timperley (2007), in order to form effective feedback, teachers should provide "clear, purposeful, meaningful, and compatible" feedback that is related learners' prerequisites (p.102).

Additionally, Supportive/constructive feedback is regarded by many researchers as an extremely important way to help learners' to be self-regulated through adjusting their own behaviour and manage the improvement of their learning process (Bose & Rengel, 2009; Carless et al., 2011; Nicol & Macfarlane-Dick, 2006).

1.6 The Performance Based Assessment

Throughout the history of language acquisition and learning, there has been always a link between the language acquired knowledge, and its actual use. Many educators and linguists have described this strong relationship differently, such as: content and process, competence and performance, langue and parole. However, they have all shared the same basic view that counterbalanced between the importance of both aspects in language learning

(Resnick & Klopfer, 1989). Therefore, in order to assess and test the language learners, it is essential to take these two aspects into consideration.

In this context, there has been a consensus among cognitive psychologists, who are mainly specialised in language learning, about the fact that:

Mere acquisition of knowledge and skills does not make people into competent thinkers or problem solvers. To know something is not just to passively receive information, but to interpret it and incorporate it; meaningful learning is reflective, constructive and self-regulated (Bransford & Vye, 1989; Wittrock, 1991; Marzano et al., 1988; Davis et al., 1990; as cited in Herman, 1992, p. 15)

Additionally, researchers and scholars have assumed that the PBA supports not just the learning process, but also affects the instructional practices positively by providing valuable data to the instructors (SRN Informational booklet. 2008).

The terms: alternative assessment, authentic assessment, and performance-based assessment are used interchangeably and synonymously. Accordingly, Herman et al. (1992) suggested that these terms similarly present any type of assessment tools that lead learners to create and produce instead of selecting the right answer. It demands not just retrieving the information from the courses, but also the use of the learners' prerequisites as well as the target skills, for the accomplishment of different authentic tasks in problem-solving situations.

In fact, the term alternative has been linked to the PBA since its very first appearance because it has come to alternate and replace the traditional types of assessment. The term Alternative Assessment has been defined, by the Longman Dictionary of Language Teaching and Applied Linguistics (Richards and Schmidt, 2002), as:

...various types of assessment procedures that are seen as alternatives or complements to traditional standardized testing. Traditional modes of assessment are thought not to capture important information about test takers' abilities in a L2 and are also not thought to reflect real-life conditions. Procedures used in alternative assessment include self-assessment, peer assessment, portfolios, learner diaries or journals, student-teacher conferences, interviews, and observation. (p. 23)

1.6.1 Definitions

The performance-based assessment or the performance assessment is an ongoing process of measuring learners' communicative skills, strategies, and language abilities to perform and use the target language in authentic tasks and real-world situations. It requires learners to produce, generate, and create rather than retrieve abstract knowledge (Rudner & Boston, 1994; Wiggins, 1989).

According to the Longman Dictionary of Language Teaching and Applied Linguistics (Richards and Schmidt, 2002), performance-based assessment is, "An approach to assessment that seeks to measure student learning based on how well the learner can perform on a practical real task" (p. 392).

The PBA is characterised with the use of interactive tasks that create authentic situations and interactive tasks to the language learners to actually perform the intended behaviour to be measured. In this respect, Brown (2004) indicated that testing learners should measure their performance in different levels including: their speaking, listening, reading or writing, and in different communicative situations as in: requesting, agreeing, disagreeing, or responding to different integrated situations. However, traditional Paper-pen tests undoubtedly cannot demonstrate and evaluate such communicative performance.

Rudner and Boston (1994) defined The PBA as a process of evaluation whereby learners are involved actively in authentic tasks, that demand producing and creating responses or products, in order to expose their skills, knowledge, and learning strategies. They added that, if performance based assessments are applied in schools, they can bring significant development to the learning process in that they help to assess learners' skills and abilities to do things instead of just evaluating what they know. i.e. the PBA is a set of meaningful, authentic, and engaging tests that aims to assess learners' skills and abilities to use the acquired knowledge. Additionally, it has a positive impact on both, the instructional practice and the learning process (VanTassel, 2013). Similarly, Stiggins (2001, as cited in Espinosa, 2015) stated that:

“Performance-based assessments represent a set of strategies that asks students to use their knowledge and skills to create a product or perform a task that is authentic and meaningful to them based on certain predetermined criteria. For instance, the tasks may include giving oral reports, writing essays, participating in interviews, among others...It involves students directly and deeply in their own learning process, and increases their confidence and motivation to learn since it emphasizes progress and achievement rather than failure and defeat (p. 2442)

Unlike the traditional paper pen tests, the PBA stimulates the ability of the learner to use the acquired knowledge, and skills of the target language so as to perform in a real world situations and authentic tasks. In the same vein, Weigle (2002) affirmed that, “The term performance assessment is used to describe any assessment procedures that involve either the observation of behaviour in real world situation or the simulation of real word activity” (p. 46).

1.6.2 Authenticity in the PBA

In EFL classrooms, the PBA is designed in forms of authentic tasks through which the learners can demonstrate their actual level where they are involved in real world tasks and authentic situations to use the target language actively, use their creative thinking, and their problem solving abilities (Espinosa, 2015).

Authenticity in assessment practices in EFL classrooms creates realistic situations that give learners opportunities to reveal their skills and abilities in using the language appropriately through solving ambiguous problem situations, use their critical thinking and creativity, as well as their communicative skills. As it was state by Wiggins (1989): An authentic test enables us to watch a learner pose, tackle, and solve slightly ambiguous problems. It allows us to watch students marshal evidence, arrange arguments, and take purposeful action to address the problems (p.705).

For ensuring the standard of “Authenticity” in the PBA, The National Capital Language Resource Centre (NCLRC, 2011, as cited in Lim & Griffith, 2011) has presented a set of criteria that should be followed while designing tasks of this type of assessment (p.3):

- Be built around topics or issues of interest to the students;
- Replicate real-world communication contexts and situations;
- Involve real problems that require creative use of language rather than simple repetition;
- Require learners to produce a quality product or performance;
- Evaluation criteria and standards are known to the students;
- Involve interaction between assessor (instructor, peers, self) and person assessed;
- Allow for self-evaluation and self-correction.

- Serves as a tool to diagnose learning and teaching problems, and has as its primary focus the continuous improvement of learning for all students (Bejarano & Gordon, 2009).

1.6.3 Tasks of the Performance-Based Assessment

Previous research on assessment has developed and presented various and distinct performance-based assessment tasks for language educators used for different purposes. Ellen (2002, as cited in SOTO et.al, 2017) proposed a list of BPA tasks and classroom activities summarised as follows:

- Constructed response.
- Cooperative–group work.
- Display, Exhibitions, Class presentation.
- Problem-solving.
- Student conference/oral interview/ Debate team.
- Group investigation
- Role-play

Some other types of PBA practices and tasks were proposed by other researchers including: (Rudner and Boston, 1989; Hoerr, 2000, as cited in Cheriet and Hoadjli, 2022, p. 6-7):

Project work is a cooperative task that assists learners not just to demonstrate what they know, but also to reflect both their personal and interpersonal skills in order to make discuss and clarify their thoughts. According to Hoerr (2000), project works can provide opportunities to work and interact in pairs and groups where learners can:

- Improve their communicative skills
- Promote learning autonomy

- Develop Social and emotional intelligence
- Develop group responsibility

The second category include Essays. Using essays and all different types of written compositions, such as: analytical, explanatory, and descriptive writings or summaries, in the PBA, plays a crucial role in demonstrating learners' effective use of the acquired knowledge. Additionally, these written compositions can be useful tools to assess learners' productive skills, including the correct spelling, language structures, and grammar (*Rudner & Boston, 1994; Wiggins, 1989*).

In EFL classrooms, a Portfolio is a form of PBA used mainly as a formative assessment. It is a collection of the learner's individual and personal works and productions (such as: essays, lab reports, classroom written tasks, and activities or reading) gathered within one file or a folder to present a full vision of the learner's performance progress throughout a given period of time. It provides an opportunity for the learners to evaluate their own improvement (*Rudner & Boston, 1994; Wiggins, 1989*).

Hoerr (2000) describes portfolios as: Keeping a portfolio for each child—a collection of work and artefacts that give a picture of the child's growth—is a way of capturing progress without using paper and pencil measures (p.27).

Since performance based assessments' tasks are formative in nature, they can be used as tasks for a formative assessment to assess EFL learners' performance and skills in higher education. This approach to assessment is also referred to as assessment for learning (**AFL**).

1.6.4 The Performance-based assessment variation

Recently, a new variation to the performance-based assessment has been created to be applied in foreign language classrooms. This innovative assessment model, the so-called:

Integrated Performance Assessment, can be a new path that leads to learners' skills development. It aims to demonstrate what foreign language learners could do with the target language instead of testing what they know (Schenck, 2019). Respectively, in the present study, the researcher opted for the Integrated Performance assessment model, as an alternative to the traditional assessment practices at the Algerian university level to be implemented as an assessment for learning.

1.7 The Integrated Performance Assessment as an Alternative

Integrated Performance Assessment (IPA), as a variation of PBA, was defined by Brown (2004) as a classroom-based performance assessment model that provides a real-world context for assessing real-world competence in the five goal areas (Which are much broader than just linguistic proficiency). It requires the learner to complete an authentic task in a specific theme (Adair-Hauck et al, 2014).

It has gained considerable acknowledgement since its first proposition by the American Council on the Teaching of Foreign Languages (ACTFL). It has received a U.S. Department of Education International Research and Studies grant in 1997 to design an assessment prototype in response to the prevailing detachment between assessment research and practice, as well as to address the need for a way to assess learner progress in attaining the standards (Glisan et al., 2006).

According to the same source, this model can be used to evaluate learners' communication skills in the three modes of communication (Interpretive, Presentational, and Interpersonal) identified in the World-Readiness Standards for Learning Languages (2015). The World-Readiness Standards incorporate the Standards for Foreign Language Learning: Preparing for the 21st Century (1996), and Standards for Foreign Language Learning in the 21st Century (2006) (Adair-Hauck et al., 2014).

The IPA model also provides opportunities for learners to demonstrate their ability to communicate around a specific theme across these three modes of communication while constantly having the opportunity to receive feedback in order to improve their proficiency levels. Core features of the IPA include: its emphasis on a “cyclical approach” to assessment, which combines continuous modelling, practice, performance, and feedback, its continuous feedback provision to learners throughout the assessment process, and the use of authentic texts for the interpretive task (Davin et al., 2011).

Several other previous studies highlighted the benefits of IPA implementation in EFL classrooms. Their results have confirmed that this type of performance assessment can help to improve learners' language communicative skills and proficiency, serve as a consciousness-raising technique to encourage teachers modifying or refocusing their instructional plans to better meet their learners' needs. Hence, it can have a positive “washback effect” on both learning and instruction (Glisan et al.; 2003)

1.8 Assessment in Higher Education

Since The main goal of teaching EFL in higher education lies on developing and enhancing students' language proficiency and skills, educators and researchers have emphasised on the importance of assessment as crucial pedagogic tool for profound learning (Murray, 2019). Accordingly, assessment can play a crucial role in evaluating and developing learners' proficiency including their learning process, knowledge and skills in order prepare them to use the language appropriately outside the classroom boundaries. In this regard, Masuku et al. (2020) specified that “Assessment is considered a crucial pedagogical instrument to enhance quality learning and to bring out qualities that are essential to preparing learners for success and becoming socially responsible citizens” (p.276).

Over the last decade, research on assessment has presented various approaches to assessment that has assisted to enhance the learning process. Among these approaches: formative assessment, assessment of learning, assessment as learning, assessment for learning and learning-oriented assessment. Each one of them has its own features and effective uses. Therefore, the selection of appropriate assessment among these methods is one of the most significant teachers' roles. Well-selected assessment can affect learning positively and also lead learners to improvement, however, flawed assessment can impede and mislead the learning process (Masuku et al., 2020).

1.8.1 Assessment in the Algerian Higher Education

Since the academic year 2003-2004, the Algerian universities have launched the LMD (License- Master-Doctorate) system, which is based on communicative approach. It has been implemented to meet the international educational standards and cope with the world's technological, and economic advancement. Also, among its main aims is to prepare and improve learners' potentials and skills and enhance their performance to what suits the international market demands and interests by providing academic and professional degrees for greater opportunities for employability.

The implementation of the LMD system was accompanied with an entire reshaping of the educational system as a whole in higher education including the remarkable change in the assessment practices and evaluation techniques. This change was the shift from summative assessment, which was the main measurement tool for evaluating the students and giving judgments, into a combination of both summative, as a final examination, and formative assessment, as an ongoing process, conducted on two separate semesters.

This was clearly stated in the practical guide of the implementation and the monitoring of the LMD system (Decree N° 137) dated back to 20/06/2009, which has included that the

learners' skills and knowledge are assessed through a combination of the two forms of assessment, a final examination, and regular continuous control. However, priority has to be placed on continuous monitoring (Sarnou & al., 2012).

At the Algerian universities specifically, current assessment practices seem to neglect the significance of using assessment to help learners in improving their language learning process. Generally, they impose a combination of two types of assessment practices, which are the formative and summative assessments, as complementary ways of assessing EFL students. However, the selection of assessment tasks mainly focuses on measuring and grading learners' language accuracy, grammar, and course contents, including taking by heart what has been learned in history and literature. Thus, there is no specific course that focuses on establishing the links between abstract knowledge and the learners' performance, as the actual usage of what has been learned.

Accordingly, it is becoming inevitable to put into practice new alternative assessment practices that would put emphasis on what learners can do instead of measuring what they know. Also, there is a strong need for an alternative assessment that considers students' communicative skills and helps to develop their language proficiency. In this respect, the Integrated performance assessment, as a highly recommended approach to assessment in higher education, can be the best choice to fill in this gap in the Algerian current assessment system.

Conclusion

This chapter presented a general literature review about assessment in general and the PBA in particular. Then, it spotted light on the Integrated-performance Assessment Model, which is a variation of the PBA, that can serve as an assessment for learning, to be proposed as an alternative in the Algerian EFL classrooms in higher education. At the outset, the researcher came to clarify and assert certain points. First, throughout the history of language teaching and learning, assessment has been so long regarded as an integral part of instruction. It has captivated the interests of educators and researchers and still playing a crucial role in language classrooms until the present day. Assessment has also passed through different stages and knew a continuous progress. It has expanded all over the world and then shifted from traditional forms into the new alternative authentic forms of the BPA assessment to cope with all the educational reforms.

Further, this chapter discussed how many assessment specialists and researchers have proved the effectiveness of the implementation of the PBA as an alternative for the last decades in enhancing the learners' skills and competencies and the learning process as a whole. Then, it sheds light on the IPA as an alternative approach to traditional assessment practices. As a variation of the PBA, this assessment model can be one of the most effective assessment models which is built upon assessing learners' abilities to use the English Language, in real contexts and situations. Additionally, the IPA can be used as a form of assessment for learning to help learners enhance their language skills and their learning process.

Finally, the end of this chapter sheds light on the gap in the assessment practices in Algerian higher education within the LMD system. Mainly, assessment in this context is a combination of formative and summative assessment. The latter is generally based on paper-

pen examinations and presentations, as the major tools to evaluate EFL learners, whereas the formative assessment is a process of measuring learners' abilities and performance along with the teaching\learning process using various PBA tasks. In spite that the PBA tasks are effective and authentic, however, they seem to be not sufficient enough to determine if students know how to do things with the acquired language or not. Simply put, because the main focus of these tests is to measure to what extent the learning objectives are attained and whether the contents of the courses are delivered in the forms of abstract knowledge, not take into account the assessment or enhancement of the learners' language skills. Therefore, the end of this chapter suggested the IPA model as an alternative to traditional assessment practices in higher education at the Algerian universities.

Chapter Two:

The Integrated Performance Assessment and EFL Proficiency

Introduction

This chapter starts first by presenting the Integrated Performance Assessment model as the first variable of the study. It provides a Review of Literature on the IPA and its use in language classrooms. Then, offers a brief description on the nature of the Integrated Performance Assessment (IPA) and its tasks. After that, it gives an overview on its use, its three modes of communication, and its implementation in Foreign Language classrooms. Additionally, it provides an explanation for both the co-constructed feedback of the (IPA) and the scoring rubrics of each phase. Moreover, this chapter presents the second variable of this study that is learners' language proficiency. Further, it explains how to develop learners' language proficiency through assessment practices. Finally, the end of this chapter highlights the importance of implementing the IPA in Language Classrooms in enhancing and assessing learners' language proficiency.

2.1 Integrated Performance Assessment

During the last decade, assessment has witnessed a remarkable shift from the traditional towards a more authentic paradigm. However, no specific assessment was devoted to measuring and at the same time improving the learners' proficiency in regard to the standards (Adair-Hauck et al., 2013). Thus, researchers and instructors claimed that although assessment has become more authentic and standard-based, it needs to be educative and pedagogic in that it connects instruction with evaluation and it should have a positive effect on both teaching and learning (Adair-Hauck et al., 2013; Wiggins,1998; Ching-Hsuan et al.,2018). In this regard, Wiggins (1998) Indicated that, "The aim of assessment is primarily to educate and improve student performance, not merely to audit it" (p. 7).

As a response to this growing need, the American Council on the Teaching of Foreign Languages (ACTFL), which is an American organisation that aims to enhance and expand

foreign language learning and teaching all around the world (Davin et al.,2011), initiated the project of creating a new assessment instrument known as The Integrated Performance Assessment (IPA).

This project was supported by the U.S. Department of Education International Research and Studies Program Award, which took full responsibility for developing its rubrics and testing its validity, reliability, and feasibility (CARLA, 2019). Their innovative Assessment model was developed to assess learners' proficiency in foreign language learning and to enhance their language communicative skills.

Among the pioneering researchers and creators of the IPA are: Eileen Glisan, Bonnie Adair-Hauck, Keiko Koda, Elvira Swender, and Paul Sandrock, in addition to hundreds of language researchers and educators, who took part in developing the first IPA model across the USA. (Davin et al.,2011).

Its early beginnings were between the late 1990s and early 2000s as a result of the need to assess the learners' performance and skills in foreign languages regarding the world's communication standards (Zapata, 2016). It was developed, based on a sociocultural perspective, unlike the traditional assessment that was constructed on a cognitive perspective (Donato, 2000). It shifted assessment practices from a traditional memorisation-based to a performance and communicative-based assessment that provided more interaction between the instructor, and the learners, or between peers in EFL classrooms (Madison, 2019).

The ACTFL's initial work was funded by the U.S. Department of Education International Research and Studies in 1997. The project was managed by approximately 40 language instructors (Adair-Hauck et al, 2006). It also involved 1,000 Chinese, French, German, Italian, Latin, and Spanish students (Ching-Hsuan et al., 2018). Along a three-year

project work, they have developed a prototype of the IPA, and conducted widespread field tests in six different research sites. In their work, Adair-Hauck et al, (2006) aimed to:

- design an assessment modal that would be used as an instrument to measure learners' progress in Foreign Language Learning in meeting the Standards (the *SFLL* of the 21st Century);
- assess the effectiveness and the feasibility of this newly developed standard-based assessment in language classrooms;
- use the assessment prototype as an incentive for curricular and pedagogical reform and
- understand to what extent the IPA could motivate teachers to change their teaching practices to create a more effective authentic learning environment (Adair-Hauck et al; 2006, pp. 363–364).

Additionally, the IPA designers claim that the IPA project, as an innovative framework for development, can offer teachers an opportunity to reconsider their all classroom assessment practices (Adair-Hauck et al; 2006). Equipped with this framework, instructional conversations between and among learners and teachers are transformed into an ongoing dialogue regarding learners' performance (Adair-Hauck et al, 2006). Correspondingly, this study also has reflected the potential of the IPA for the washback effect as they have tracked the impact of using it on teachers' goal setting and planning for learning.

After the publication of the IPA Manuel by the ACTFL in 2003, two major studies were presented on the use of the IPA. The first was a descriptive study by Adair-Hauck et al. (2006) on the development of the IPA model in the field. Throughout this study, the researchers described the impact of the IPA on the teaching process during the field trials that were carried out in middle and high schools of FL Spanish classrooms. This study concluded that the IPA could help instructors to modify and reconstruct classroom tasks and activities to

promote standards-based curricula. Additionally, they suggested further studies and research on the implementation of the IPA at all other levels.

The second study was conducted in the same year (2013), it was presented by Bonnie Adair-Hauck, and Francis J. Troyan "A Descriptive and Co-Constructive Approach to Integrated Performance Assessment Feedback". This study was based on a critical discourse analysis and it was devoted to highlight the importance of feedback in the IPA model as well as to describe the characteristics of effective feedback. Therefore, this work now is regarded as a guideline for the implementation of the IPA in language classrooms (Adair-Hauck et al., 2003; Ching-Hsuan, 2018; Zapata, 2016).

Several studies focused on the application of the IPA in many different educational levels and environments. Davin, et al., (2011) described a project in which the IPA was implemented in fourth and fifth-grade classes with learners who had studied Spanish for at least four years. Thirty students participated in the 12-week instructional unit and end-of-unit IPA on the Caribbean Sea. The highest levels of performance were noted in the interpersonal mode and the lowest in the interpretive. It has been also found that young learners held positive attitudes towards collaborative work on the IPA tasks. According to the researchers, this was an indication of their understanding of the connection between their learning and assessment.

The IPA has also been integrated into L2 instruction at the high school level. Troyan (2007) has provided a full description of the institutional and pedagogical stages, as well as processes involved in the implementation of IPA in this level. This study also highlighted the importance of professional development, teachers' collaboration, students' understanding of the three communicative modes, and the learning–assessment–feedback cycle functions.

Another study has been conducted on the implementation of the IPA at the post-secondary level. This study was piloted by Glisan, Uribe, and Adair-Hauck in 2007. It involved an advanced class of L2 Spanish learners in Latin American culture and civilization. They examined the role of IPA in providing information about the 25 participants' performance and usage of the target language in the three modes of communication. Additionally, they have tested the feasibility of the IPA in advanced classrooms.

According to this study, learners' performance in the interpersonal mode was not as good as that of the presentational. In the latter learners perform and present better. However, in both modes (the interpersonal and presentational), learners have met the expectations. The only mode which indicated that some learners' performance was not sufficient enough was the interpretive mode because they lack listening strategies (Adair-Hauck et al., 2013).

Similar to the study of Davin et al., (2011), lower levels of performance in the interpretive mode among the participants were marked. However, unlike the study of Davin et al., the participants performed best in the presentational mode.

From all this reviewed literature, it seems that in spite that the previous studies and research on the IPA have provided valuable information and data about the implementation of this new assessment model at different levels (the novice, intermediate, and even advanced levels), however, only A few studies at the university level were conducted. This was confirmed and explained by Troyan (2012) and Zapata (2016) pointing to the fact that more research is needed in that area of interest. Moreover, another lack of studies about IPA in the Algerian EFL context was also noticed given the shortage and rare studies in this field.

2.2 The Nature of the Integrated Performance Assessment (IPA)

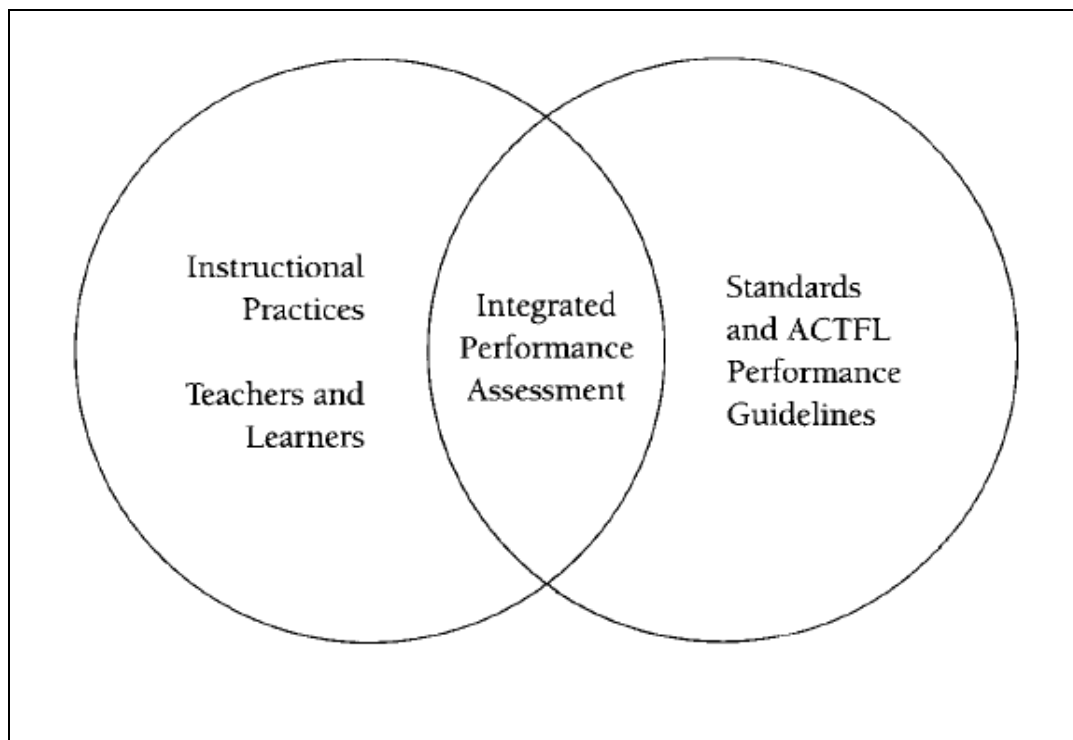
In EFL classrooms, the IPA is a type of standards-based performance assessment that has been developed to assess FL learners' language proficiency in three modes of

communication: Interpretive, Interpersonal, and Presentational based on the world readiness standards *for Learning Languages* in the 21st Century (Davin et al.,2011).

It was defined in the manual as “A multi-task or cluster assessment featuring three tasks, each of which reflects one of the three modes of communication— interpersonal, interpretive, and presentational” (Adair-Hauck et al., 2013, p.34)

Figure 1

Integrated Performance Assessment (IPA): Connecting Instructional Practices, Standards, and ACTFL Performance Guidelines,



(Source: Adair-Hauck et al., (2006) *The Integrated Performance Assessment (IPA): Connecting Assessment to Instruction and Learning*: p. 365)

According to Adair-Hauck et al., the IPA is a multi-task assessment model. It is divided into three phases based on the three modes of communication and designed within one single thematic context (Adair-Hauck et al., 2006). It tends to link between classroom instruction and learning as well as to integrate the world readiness standards, *for Learning Languages* in

the 21st Century, in language classrooms. Additionally, the IPA aims to evaluate the learners' abilities and skills in the three modes of communication, interpersonal, interpretive, and presentational (Davin et al., 2011). The aims of each phase are:

- The interpretive task aims to assess learners' ability to comprehend authentic texts literally and interpretively. Throughout this phase the teacher provides the learners with "*content knowledge*" to be acquired and then applied in the next stage.
- The interpersonal task aims at assessing learners' ability to engage in a spontaneous, negotiated conversation, to interact and to present their points of view.
- The presentational task aims at assessing learners' skills and abilities to deliver a message to audience by writing or speaking a product (Adair-Hauck et al., 2013; National Standards Collaborative Board, 2015). It provides opportunity to the learner to create, express and present.

The Centre for Advanced Research on Language Acquisition (CARLA) has defined The IPA as: A cluster assessment featuring three tasks, each of which reflects one of the three modes of communication, Interpretive, Interpersonal, and Presentational, aligned within a single theme or content area, reflecting the manner in which students naturally acquire and use the language in the real world or the classroom. (University of Minnesota, 2019)

In each task; the teacher provides a range of information and elicitations on the target topic that would help learners to complete the following task. This type of assessment can be designed differently to assess language learners at different levels of language proficiency, whether for novice, intermediate, or advanced.

According to other previous research on the IPA (CARLA, 2019; Adair-Hauck et al. 2013, Adair-Hauck, et al., 2013; Gisan et al., 2007; Diaz Maggioli, 2020) this type of assessment is:

- A standards-based, or outcome-based, which is designed with specific scoring rubrics to assess whether the learners' performance meets the expected criteria of the task,
- Performance based, that aims at measuring the learners' abilities to apply their knowledge and skills of the target topic in an integrative task.
- Developmental in nature, that it identifies the learners' strengths and weaknesses in different areas including communication, behaviours, interaction, and adaptive skills, then it helps them to improve their performance through constant feedback.
- Integrative, in that it integrates the learners in real-world situations which require them to integrate their language skills and knowledge to solve them. Also, it integrates each communicative task of the three phases with the following task.

In addition to these features of this innovative assessment model, it has other characteristics related to its original framework including the Performance-Based Assessment (PBA) goals and the Dynamic Assessment (DA) (Madison, 2019).

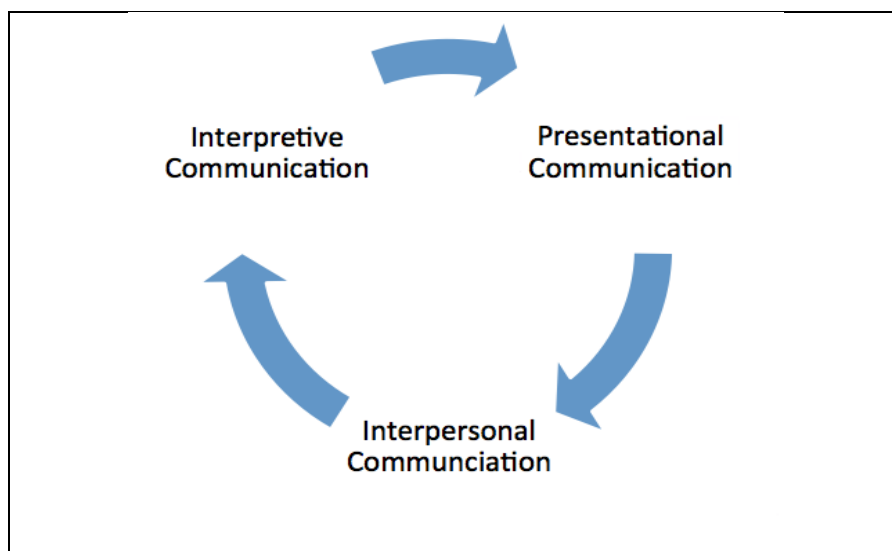
- **The PBA goals:** As it has been discussed in the previous chapter, the PBA bridges the gap between classroom instruction and real-world usage of the language through authentic situations. Therefore, the IPA requires performance-based tasks which simulate a real-world situation and requires

full integration of the learners' language skills and abilities including their lexical and cultural knowledge (Adair-Hauck et al., 2006).

- **The DA Goals:** the DA equals the importance of instruction and assessment. It reflects how can learners achieve better results and develop their abilities with the guidance of an adult. Generally, the DA requires three phases through which (1) learners are given an assessment, (2) subsequent feedback, then (3) an additional assessment. The constructive feedback provided by the instructor in the second phase plays an important role in developing the learners' competence in the target language (Antón, 2009). Based on this principle; The IPAs' feedback provided by the teacher came to enhance the learners' performance in a cyclical approach (Davin, 2013) and it is a part of the learning process itself. (Adair-Hauck et al., 2013).

Figure 2

Integrated Performance Assessment Units: A Cyclical Approach



(Source: Centre for Advanced Research on Language Acquisition CARLA, 2019)

The IPA model is designed upon a cyclical approach that helps in developing the topic and the performance of the assessment in each unit. All the units are developed around one

selected theme and divided into three main tasks that correspond to the three modes of communication (interpretive, interpersonal, and presentational) (Tedick et al., 2006)

2.3 Tasks of the IPA

The IPA is a multi-task assessment model. As it has been mentioned earlier, this assessment model is divided into three phases according to the three modes of communication (interpretive, interpersonal, and presentational), and designed within one single thematic context. Sequentially, learners start first with an interpretive phase where they complete some tasks to learn about the target topic. Then, they use the information that has been learnt from the previous tasks in the interpersonal phase in the form of productive tasks. Finally; in the third phase learners are required to summarise all what has been learnt in the presentational phase. Throughout these phases, the teacher uses clear scoring rubrics and for guiding the learners to accomplish the tasks (Adair-Hauck et al., 2006).

When conducting an IPA, the teacher has to initiate the IPA with a general introduction and presentation of the whole process and its purposes. Learners should have a clear overview about all the phases (interpretive, interpersonal, and presentational communication) and all the series of authentic tasks, that will be conducted including the theme, the context, the tasks, and so forth. (Adair-Hauck et al., 2006). Tasks included in the IPA should suit the learners' proficiency levels, and the theme has to be selected according to their interests (Adair-Hauck et al., 2013).

2.3.1 The Interpretive Mode

According to the IPA Manual (2013), the interpretive mode targets the learners' receptive skills. It includes reading, listening, viewing, or watching to range a global understanding of the target theme through inferring and Hypothesising. Moreover, the interpretive tasks are regarded as a source of information for the target theme. This phase

provides the learners with tools and data to complete the subsequent task. It affords authentic materials that expose the learners to the target language, and culture as much as possible.

Designers affirm that the interpretive mode is classified as the first task in the IPA as it provides learners with sufficient knowledge for the accomplishment of the following tasks. It also guides learners and helps them in understanding the interpretation of the target culture. Correspondingly, these learners can connect their background knowledge with the newly learnt items to provide a complete comprehension of the target theme.

The Interpretive Mode of Communication is a constructive process that provides the learners with authentic materials on the target language. It entails both literal and deep comprehension since it leads learners to make interpretations and inferences to evoke learners' interpretive skills (Adair-Hauk et al., 2016; CARLA, 2019).

Throughout this phase, learners are required to express their comprehension of the used materials at two levels: literally and interpretively. First, on a Literal level, through demonstrating their understanding of the surface meaning, or what has been presented explicitly (by words and phrases recognition and understanding meanings, Main idea detection, Supporting idea detection, Organisational principles). Then, on an interpretive level; through demonstrating their understanding of the implicit meaning. In other words, "reading between the lines" where learners involve their background knowledge and cultural understanding to present a deeper interpretation of the text (by understanding the appropriate meanings of the context and organizational features, inferring implicit ideas, Inferring the authors ideological and/ or cultural reference and insights) (Adair-Hauk, 2016, p. 90).

For authentic materials used in the interpretive task, the Manual (2016) indicated that they should be genuine documents produced by native speakers and for native speakers (Galloway, 1998). Among these materials that can be used in the interpretive mode:

- **Authentic Texts for Listening/Viewing:**

- ✓ Movie scenes, Short movies, Podcasts, Radio Broadcasts, Announcements, Television Programs and Talk Shows, Plays, Music, Art, Debates and interviews, Lectures, and Sports.

- **Authentic Texts for Reading:**

- ✓ Newspapers, Magazines, Internet Sites, Articles, Schedules, Instructions, Maps, Menus, Recipes, Text Messages, Letters, Emails, Lyrics, Poetry, Short Stories, Stories, Novels, and Plays (CARLA, 2019).

Furthermore, Adair-Hauck et al., (2013, pp. 34-36) also presented some examples of authentic resources that can be used in the interpretive phase including:

- “personal letters or email correspondence;
- Simple biographies or descriptions of people from popular culture magazines or newspapers;
- Product commercials in the target language from newspapers or magazines;
- Public service announcements in magazines and newspapers...;
- Product advertisements or sales advertisements from a supermarket;
- Interviews or surveys from youth-oriented magazines;
- Short stories;
- Advice columns of personal interest to students;
- Photo stories with captions such as the “fotonovelas”;
- Essays or editorials in authentic target culture newspapers;
- Authentic songs or poetry by artists of the target culture;
- Comic strips.”

2.3.2 Interpersonal Task

The Interpersonal Mode of Communication, as a complementary phase to the interpretive mode of communication, is a two-way interactive process that is conducted through an oral interaction only. It includes two important features that control the interactive task design Unrehearsed and Negotiated. The first means that the learners should improvise without any note cards or memorisation as they are involved in conversations or class discussions. The second characteristic means that the learners should negotiate meanings, exchange information, listen and interact with one another, respond to others' comments, and express their views and feelings about the target theme (Adair-Hauck et al., 2006; CARLA, 2019).

Moreover, according to the manual (2013), the Interpersonal mode of communication involves the learners in conversations and discussions to negotiate meanings and interact with each other. This requires them to ask for clarification, repetition, and confirmation, whenever they feel a lack of understanding.

Before starting the interpersonal communication, the instructor may need to present a model for the required skills and criteria to make the learners aware of the expected performance. Accordingly, this would help these learners to improve their performance. All this phase is advised to be recorded either by audio or videotaped for the feedback loop (Davin et al., 2011)

Subsequently, after the interpersonal task is accomplished, the instructor scores the learners using the IPA rubrics, then provides learners with sufficient feedback about their performance (Adair-Hauck et al., 2006).

According to Adair-Hauck et al. (2006) "This type of performance-based feedback will assist the learner to self-assess and self-adjust his/her language Performance" (p.370). In the

same vein, Wiggins (2004) noted that “The more self-evident the feedback to the performer; the more likely the gain” (p. 5). Therefore, the feedback is very essential not just during this communicative mode but in the IPA as a whole to enhance the learners’ performance to meet the standards.

2.3.3 The Presentational Mode

The Presentational Mode is a productive process that involves learners’ productive skills. It entails either speaking, writing, or both throughout productive and presentational tasks by using the information attained from the preceding tasks. The learners’ products, in the form of a presentation, a speech, an interview, a video, a report, a story, or a newscast, are used for communication to be addressed to listeners or readers. These products represent an authentic situation that simulates a real-world usage of the target language (Adair-Hauck et al., 2013).

The improvement of learners’ performance during this phase is based on two aspects. Firstly, before starting the work on the presentational Task, the instructor should present scoring rubrics to the learners to make them aware of the criteria, and the qualities expected to achieve an excellent product (Adair-Hauck et al., 2013). Secondly, once the presentational task is accomplished by learners and evaluated, the instructor should provide ultimate feedback on this phase (Adair-Hauck et al., 2006).

The Presentational mode is vital to assess learners’ communicative skills and enhance their performance and communicative skills through sequential feedback. Therefore, it may include several drafts in the written communication or several rehearsals of oral presentations. Unlike the interpersonal mode, it is permitted for learners to use pre-planned talk or note cards during the oral rehearsals. While drafting, the teachers’ feedback is very important to guide the learners. Through feedback, learners can correct their errors and then enhance their

performance and content for a final high-quality presentation. Thus, helpful feedback can be received, not just from the teacher, but also from classmates, friends, or family provided that they have previous knowledge about the task rubrics and requirements (Maggioli, 2020).

To have good results in the presentational phase, Adair-Hauck et al. (2013, p. 37) recommend in the Manual some suggested strategies to prepare and warm up students for the task:

- “beginning with warm-up activities that lower the affective filter and provide students with plenty of planning time;
- incorporating a process-oriented approach to presentational tasks with phases for drafting, peer editing, revising, and rewriting;
- offering feedback to students that includes attention to the message itself in addition to linguistic accuracy.
- providing periodic opportunities for students to share their work with audiences other than the teacher and receive feedback from them; and
- periodic videotaping of students' presentations and having students analyse their own work.”

2.4 The Three Modes of Communication

For better achievements, researchers (Adair-Hauck et al. 2013a, Adair-Hauck, et al., 2013b; Gisan et al., 2007; Maggioli, 2020) asserted that it is vital to switch between the interpretive and interpersonal modes starting with the interpretive if the teacher finds that the learners are not ready to interact spontaneously in an oral discussion or starting by the presentational tasks before interpersonal (CARLA UMN, 2016). According to the Centre for Advanced Research on Language Acquisition (CARLA, 2019), this swap

between the phases will provide learners with adequate time and opportunities to practise the new vocabulary, structures, and information that have been acquired in the interpretive phase as well as constant feedback on how to use them appropriately to communicate.

In short, the ACTFL has summarised all the descriptions of the three modes of communication in table 1.

Table 1
The IPA's Three Modes of Communication

Interpretive	Interpersonal	Presentational
<ul style="list-style-type: none"> • Interpretation of what the author, speaker, or producer wants the receiver of the message to understand 	<ul style="list-style-type: none"> • Active negotiation of meaning among individuals 	<ul style="list-style-type: none"> • Creation of messages either oral or written
<ul style="list-style-type: none"> • One-way communication with no recourse to the active negotiation of meaning with the writer, speaker, or producer 	<ul style="list-style-type: none"> • Participants observe and monitor one another to see how their meanings and intentions are being communicated 	<ul style="list-style-type: none"> • One-way communication intended to facilitate interpretation by members of the other culture where no direct opportunity for the active negotiation of meaning between members of the two cultures exists
<ul style="list-style-type: none"> • Interpretation differs from comprehension and translation in that interpretation implies the ability to read (or listen or view) "between the lines," including understanding from within the cultural mind-set or perspective 	<ul style="list-style-type: none"> • Adjustments and clarifications are made accordingly 	<ul style="list-style-type: none"> • To ensure the intended audience is successful in its interpretation, the "presenter" needs knowledge of the audience's language and culture
<ul style="list-style-type: none"> • Reading (websites, stories, articles), listening (speeches, messages, songs), or viewing (video clips) of authentic materials 	<ul style="list-style-type: none"> • Speaking and listening (conversation); reading and writing (text messages or via social media) 	<ul style="list-style-type: none"> • Writing (messages, articles, reports), speaking (telling a story, giving a speech, describing a poster), or visually representing (video or PowerPoint)

(ACTFL Performance Descriptors for Language Learners, 2012)

2.5 The co-constructed feedback of the (IPA)

One of the most significant steps that learners should go through in the IPA is the feedback phase. During this session, which comes after each phase, learners are given an opportunity to evaluate, discuss, and correct their performance flaws in the form of co-constructed feedback.

Unlike the traditional feedback that comes after the assessment process in which the teacher provides the feedback to students, co-constructed feedback is a process through which the teacher assists the learners to construct their own feedback through the use of assisting and assessing questions after each phase. That is to say, it is a collaborative conversation whereby the learner is “playing the role of a participant rather than a recipient” (Adair-Hauck et al., 2013, p. 28). It can lead learners towards proficiency and success by providing them with the expected criteria to be used for student self-assessment.

This type of feedback follows a social constructivist perspective. It was described by Adair-Hauck and Troyan (2013) as: “The dialogic interaction in which the expert (the teacher) and the apprentice (the student) co-construct a performance assessment” (Adair-Hauck et al., 2013, p. 25).

Co-constructed feedback has specific characteristics that should be taken into consideration by teachers while forming it. The following figure 3 includes a summary that presents these characteristics.

Figure 3*Characteristics of a Descriptive and Co-Constructive Approach to IPA Feedback*

1. Focus on learner performance compared to model performance.
2. Language development is a work in progress.
3. Questioning as part of the discussion.
4. Mutual appraisal of performance is integral to descriptors.
5. Shifting responsibility of learner's own role in learning/assessing.
6. Brainstorming strategies to improve performance.
7. Active participation of learners to self-assess and self-regulate.
8. Reflective process for both the teacher and the learners.
9. Lack of evaluative or judgmental statements.

(Source : Adair-Hauck et al., 2013).

Moreover, the discourse of the co-constructive feedback used in the IPA model is mainly a dialogue. It consists of assisting and assessing questions addressed by the teacher to the learners to help them to realize their flaws and compare them to the expected performance. In this way, learners will be able to enhance their performance next phase, then achieve better results. According to the Manual (2016), "Assessing questions determine the level at which the student can perform without assistance, whereas assisting questions support the student to reach a level of performance that he or she could not reach unassisted" (p. 305).

The following figure 4 provides a descriptive summary of the Discourse Features used in the IPA's Co-Constructive Feedback.

Figure 4

Discourse Features used to Facilitate a Co-Constructive Approach to IPA Feedback

1. Dialogic or symmetrical combination of teacher and student talk
2. Judicious use of explicit directives and meta-statements; mainly used initially to de
3. Use of the first person collective pronoun « we » for joint problem solving
4. Sparse use of assessment questions
5. Abundance of assisting questions and cognitive probes
6. Use of presuppositions, ellipsis, and abbreviated speech (mirroring everyday language)
7. Language, especially questions, needs to be attuned to a level where performance requires assistance

(Source: Adair-Hauck et al., (2013). p. 122)

Additionally, the designers of this innovative assessment model have offered a detailed outline suggested for teachers to help them to use the Co-constructive feedback in their classrooms. This guiding outline includes: familiarity with the tasks, effective feedback, assistance and guidance of the teacher, calling learners' attention to flaws in their performance for improvement, helping the learners to do self-assessment, and constructing their own feedback. These suggestions are explained in figure 5.

Figure 5*Teachers and Learners as Co-Constructors of IPA Feedback*

1. To set the learners up for success, ensure that they are familiar with the IPA rubrics. Providing models of student performance across the modes will enable them to better understand the question, “Where am I going?” or “What are the goals?”
2. During the feedback loop, both the teacher and learner review language performance (via audio or video if interpersonal or presentational speaking), then individually they assess the performance using the IPA rubric. Allowing the learner, the opportunity to self-assess first will shift more responsibility onto the learner. When the student is not sure how to assess his/her performance, is confused, bewildered, etc., the teacher’s assisting questions may help the learner to better understand and assess his/her performance. After the learner’s appraisal, the teacher matches his/her rating with the student. Most importantly, the teacher and learner should assess and acknowledge why the performance is at a particular level, i.e., “How am I going?”
3. Through the use of the teacher’s critical assisting questions or cognitive probes, co-constructive IPA feedback shifts or places the responsibility on the student to appraise and assess his/her performance.
4. Cognitive probes will also prompt and support the learners to perceive, observe and examine their work, that is to think about their progress in language learning.
5. The teacher and students together discuss strategies that will help to improve performance and plan for future achievement/performance, or “Where to Next?”
6. Co-constructive IPA feedback allows for both the teacher and the students to be learners in the IPA assessment/learning process.

(Source : Adair-Hauck et al., (2013). pp.132-133)

2.6 The Implementation of the IPA in Foreign Language classrooms

Adair-Hauck and her colleagues (2006), in designing the IPA, proposed a list of questions that teachers can answer with the implementation of the IPA. These questions are:

- “Am I assessing performance using standards-based and real-world tasks that are meaningful to students?”
- “Am I assessing the same way that the students are learning?”
- “Are the students able to demonstrate survival skills in the target language?”
- “How can I move beyond isolated, single-skills assessment?”
- “How can I more effectively assess the interpretive skills of my students as they relate to the ACTFL Performance Guidelines for K-12 Learners?”
- “What kind of feedback will improve learner performance?” (Adair-Hauck et al., 2006, p. 365)

The progression of tasks and Phases of this innovative assessment tool is built upon Bloom's taxonomy of the educational objectives. The latter is one of the most famous and useful theories used in the current teaching and assessment objectives. This theory aims at helping the teacher to develop learners' higher-order thinking, achieve learning objectives that concern the subject being taught, help learners to enhance their learning process as well as select accurate tests to evaluate learners' improvement (Bloom et al., 1956, as cited in Anderson, et al., 2001). Levels of Bloom's taxonomy include (Summarised from: Anderson, et al., 2001):

- Remember (Recognizing, Recalling)

- Understand (Interpreting, Classifying, Summarizing, Inferring, Comparing, Explaining)
- Apply (Executing, Implementing)
- Analyse (Differentiating, Organizing, Attributing)
- Evaluate (Checking, Critiquing)
- Create (Generating, Planning, Producing)

Similarly, the IPA was designed in a way that follows the order of this taxonomy (Povey, 2019). The following table 2 explains the educational objectives of each phase of the IPA based on Bloom's taxonomy:

Table 2

The educational objectives of the IPA based on Bloom's taxonomy.

Interpretive	<ul style="list-style-type: none"> • Students Understand authentic materials through: Interpreting, Classifying, Summarizing, Inferring, Comparing, Explaining
	<ul style="list-style-type: none"> • Feedback phase (Checking, Critiquing)
Interpersonal	<p>Students negotiate meanings and interact with each other by applying the knowledge acquired in the first phase. They observe each other's views, answers, and intentions to compare, adjust, and clarify their participation. Accordingly, and based on bloom's taxonomy, it can be said that during that phase students:</p>
	<ul style="list-style-type: none"> • Remember (Recognizing, Recalling) • Apply (Executing, Implementing) • Analyse (Differentiating, Organizing, Attributing)
	<ul style="list-style-type: none"> • Feedback phase (Checking, Critiquing)
Presentational	<ul style="list-style-type: none"> • students produce a comprehensible and accurate oral and written presentation that would be clear to the audience. Thus, students pass through the final stage in bloom's taxonomy:
	<ul style="list-style-type: none"> • Create (Generating, Planning, Producing)
	<ul style="list-style-type: none"> • Feedback phase (Checking, Critiquing)

In each phase of the three modes of communication, the IPA makes use of one of the language skills (Reading, Writing, Listening, or Speaking) to support the other skill in the next phase. Thus, each of the integrated skills helps in developing the other in a cyclical approach. The first phase, for instance, integrates receptive

skills (Reading, Listening, or both of them) to help the students to acquire a repertoire of knowledge including terms, facts, and ideas that will be used in the next phase. This acquired repertoire of knowledge plays a vital role in enhancing learners' productive skills (speaking and writing) since it allows them to use what they learned to participate and speak accurately in the interpersonal phase then write, create and present in the presentational phase (Madison, 2019).

2.7 Scoring Rubrics of the IPA

Since the IPA is a newly assessment model, the designers have provided detailed rubrics to be used for scoring all tasks across the three modes of communication. Firstly, this scoring rubric is an effective tool that helps teachers to measure and evaluate their learners in an objective and accurate manner. It tends to describe learners' expected and anticipated performance in a specific task or assessment.

It was defined as a scoring monitor used to assess the quality of output. It provides a set of specific expected criteria for a particular assignment (Stevens et al., 2005; Mertler, 2001). Rubrics are not just a kind of checklists to judge and score learners' performance but also a predesigned plan and clear description of all the criteria and the aspects expected on students' performance that positively helps in enhancing students' learning process (Panadero et al., 2013)

Jonsson and Svingby (2007) identified that rubrics can be used as a teaching tool that simplifies for the learners the way how to practise self and peer evaluation as well as the way of generating feedback.

Mainly, there are two different types of rubrics, which are holistic and analytic. A holistic rubric involves assessing the achievement and the quality of an

overall product or performance without paying attention to details or components. Whereas, the analytic rubric uses details to assist the instructor in evaluating students' achievements.

(Andrade 2000).

According to the Manual of the IPA (2016), the type of rubrics used in the IPA is both generic and analytical. Generic in that “they describe characteristics of language performance without specifying particular content or task details”; and analytic “as they include a set of criteria for performance with a range of descriptions for each criterion” (Adair-Hauck, et al. 2013. p .49).

The teacher should share rubrics with the learners because they can help to enhance the learning process. Firstly, the shared scoring rubrics provide learners with the criteria of intended and expected performance, which helps them to realise their flaws and enhance their performance. Moreover, it creates a challenging atmosphere in which learners compete to enhance their performance and arrive to the top level (Adair-Hauck, et al. 2013).

2.7.1 IPA Interpretive Rubric

Since the interpretive phase integrates only the receptive skills (Listening, Viewing, or Reading) its rubric involves evaluating: literal and interpretive comprehension.

- **Literal Comprehension entails:** Word Recognition, Main Idea Detection, and Supporting Detail Detection.

- **Interpretive Comprehension includes:** Organisational Features, Guessing Meaning from Context, and Inferences (Reading between the Lines).

Based on these criteria, the scoring rubric of the IPA model, designed for this study, was divided into two parts. Part One, that is devoted for literal comprehension, is scored on nine (09) points; whereas, Part Two, which is devoted to interpretive comprehension, is scored on eleven (11) points. More details on this rubric are provided in the Methodology Chapter.

2.7.2 IPA Interpersonal Rubrics

The rubric, that assesses the IPA interpersonal tasks, is based on evaluating the following aspects of performance defined by the Manual (2016, p.53):

- **Language Function:** language tasks the speaker is able to handle in a consistent, comfortable, sustained, and spontaneous manner; e.g., posing questions to a restaurant waiter/waitress to make a decision regarding what food to order.

- **Text Type:** Quantity and organisation of language discourse, on a continuum from words, to phrases, to sentences, to connected sentences, to paragraphs, to extended discourse.

- **Communication Strategies:** The quality of engagement and interactivity in the conversation; the amount of negotiation of meaning— i.e., strategies for clarifying meaning in the face of breakdown in communication; how one participates in the conversation and advances it.

- **Comprehensibility:** Who can understand this person's language? Can this person be understood only by sympathetic listeners used to the speech of language learners? Can a native speaker unaccustomed to non-native speech understand this speaker?

- **Language Control:** Grammatical accuracy, appropriate vocabulary, degree of fluency.

The oral responses and answers during this phase are evaluated and graded based on these criteria. According to it, the “Exceeds” performance could be assigned 4 points, the “Meets—Strong” could be assigned 3 points, the “Meets—Minimal” would carry 2 points, and the “Does Not Meet” would carry 1 point. The total score obtained from the five evaluated aspects of the language “Language Function, Text Type, Communication Strategies, Comprehensibility and Language Control” is 20 points (more details on the “Interpersonal Mode Rubric with Scoring Points” in the Methodology Chapter).

2.7.3 IPA Presentational Rubric

Similar to the interpersonal rubric, the presentational rubric includes detailed criteria of. Language Function, Impact, Communication Strategies, Comprehensibility and Language Control, however, as it can be noticed only one new aspect is added to the list which is the “Impact”. Again, according to the Manual (2016), **Impact:** is “A response from the audience of listeners/readers/viewers and/or holds their interest; i.e., in assessing impact, one might ask whether the performance is “powerful, memorable, provocative, or moving” (p. 56). Additionally, since the presentational phase involves the productive skills where students are required to write and speak other criteria should be included too such as: “fluency accuracy and form”.

2.8 The importance of implementing the IPA in Language Classrooms

The previous studies on the IPA have shed light on the impact of the IPA in Language Classrooms. They have proved that the implementation of this innovative type of assessment can bring enlightening results to language learning. According to their results, the positive impact, traced from the IPA improve, can influence both the learner and instructors.

Firstly, the IPA helps in improving learners' communicative skills by providing various opportunities for them to demonstrate their abilities across the three modes of communication and receive effective feedback in parallel (Tedick et al., 2006).

Moreover, the IPA helps to motivate language teachers to integrate the three modes of communication in their lesson plans, and to reconsider their teaching practices to create more authentic learning environments (Ching-Hsuan et al., 2018).

Additionally, Adair-Hauck et al., in their study "The Integrated Performance Assessment (IPA): Connecting Assessment to Instruction and Learning", have found that the IPA has a potential impact on classroom instruction, and learning. Among their results is that:

- The IPA, as a dynamic and an interactive assessment process, improves learners' performance as well as the whole process of teaching and learning as it bonds between: Performance-based Tasks, Authentic Material, Longitudinal Rubrics and Expectations of how/what learners are expected to perform all along with effective constant feedback;
- it tends to improve the curricula;
- it works in a cycle of linking between instruction, learning, and assessment;
- it may have a constructive "washback effect" on instruction since it motivates teachers to improve their practices and to adopt more effective teaching techniques. It also helps on enhancing the curriculum;
- it encourages learners' self-assessment, self-reflection, and self-regulation;

- it raises awareness of the different modes of communication and world-readiness language standards and how to integrate them into language teaching; and
- it provides opportunities to the learners to determine their strengths and weaknesses and to develop their communicative skills (Adair-Hauck et al., 2006 ; Adair-Hauck et al. 2013; Gisan et al., 2007 ; Madison,2019 ; Diaz Maggioli, 2020)

Furthermore, according to the Manual “Implementing the Integrated Performance Assessment” (2016), this innovative assessment model can have significant advantages on foreign language learning including:

- It provides EFL learners with a great chance to practise and reveal their actual communicate abilities and skills.
- It helps the teacher to measure and evaluate the level at which learners deal with authentic texts in the target language.
- It offers the learners opportunities to interact with each other whether in oral or written form using the target language.
- It assists EFL learners to learn how to create and present oral and written products to audiences then evaluate their works.
- It evaluates what the learners can do with the language rather than retrieving knowledge. In other words, the IPA entails tasks that tests the students' abilities to use the acquired repertoire of knowledge and skills in authentic situations rather than meaningless tasks.
- Throughout the three phases, the IPA helps learners to enhance their performance and develop their communicative skills by providing them

various opportunities to rehearsal, practice, then getting constructive feedback.

- Authentic integrated situations of the IPA play a vital role in preparing foreign language learners for real life communication, career skills and language needs for future work opportunities.

- If the materials used in the IPA tasks include authentic cultural resources, it could develop learners' cultural knowledge and rise their cultural awareness and interest.

- It develops language learners' cognitive abilities including their reasoning, creative thinking, risk-taking as well as their higher-order thinking.

Based on what has been mentioned so far, it can be said that the IPA can serve not just as an assessment tool to evaluate students' performance, but also as a pedagogical tool and source of repertoire of knowledge that play a pivotal role in helping EFL teachers to develop, and enhance their learners' language proficiency, skills, and performance.

2.9 Assessing and Enhancing EFL learners' Language Proficiency

Stern (1991) indicated that the main objective of teaching a foreign language is to assist learners to reach proficiency in the target language. Therefore, it is the teachers' role to select teaching materials and methods that can assist their learners to develop their language proficiency including suitable assessment methods. Before moving to how to develop language proficiency in EFL classrooms using assessment tasks, it is important to conceptualise and precise language proficiency meaning.

2.9.1 Language proficiency

According to the Oxford dictionary (2010) for language teaching and applied linguistics, language proficiency refers to:

The degree of skill with which a person can use a language, such as how well a person can read, write, speak, or understand language. This can be contrasted with language achievement, which describes language ability as a result of learning. Proficiency may be measured through the use of a proficiency test” (p.321).

During the last few decades, various studies have been conducted on language proficiency in the field of EFL teaching and learning. These studies aim at providing a holistic image to researchers and instructors about language proficiency. (Stephenson, et al. 2003)

Cummins (1980) referred to proficiency as the communicative language ability applied by either native speakers or foreign language speakers. He indicated that language proficiency has two dimensions:

- **Cognitive Academic Language Proficiency (C ALP):** it includes the academic knowledge that language learners acquire within a formal classroom setting. This knowledge is mainly required to achieve academic tasks.
- **Basic Interpersonal and Communicative skills (BICS):** it includes communicative skills required to communicate in common interactive situations.

Additionally, according to Canale (1983), language proficiency entails: linguistic knowledge, communicative abilities, language abilities, and language

skills that language learners acquire similar to those of native speakers. These elements in FL classrooms should be taken into consideration while planning instruction and preparing language tests that evaluate proficiency.

Furthermore, Omaggio (1986) proposed some strategies for proficiency-oriented instruction that aims at developing learners' proficiency levels. These strategies include that instructors should offer language learners sufficient opportunities to use the target language using real-life simulation tasks and contexts relevant to the target culture, integrate language learners in communicative tasks and interaction using the target language functions, consider the affective and cognitive needs of the learners to create an appropriate learning atmosphere, develop learners' cultural awareness of the target language to prepare them to be integrated into the target language community. He added that instruction should take into account developing linguistic accuracy as an initial step towards teaching a foreign language.

Stem (1997) claimed that among the most common challenges in foreign language instruction is finding methods to help language learners to enhance and attain proficiency levels. He added that proficiency in the classroom setting should be involved in a conscious process of language learning i.e. language learners have to be conscious and thoroughly aware of the importance of developing Proficiency levels in order to maintain their learning for this purpose.

Moreover, proficiency in language learning involves the effectiveness of using the four language skills: listening, speaking, reading, and writing, for diverse contexts, in a vast array of situations, with different interlocutors (TESOL, 2006). This was clearly presented in Richards and *Platt* definition (1992, p. 242) stating that proficiency refers to the ability of foreign language learners to use a language effectively in different situations in which they can comprehend the target language

through listening or reading and then communicate effectively through writing or speaking. Accordingly, Proficiency is measured by testing these language skills.

Additionally, Kern (2000) added to the conception of language proficiency three academic scopes: linguistic, cognitive, and sociocultural. He believes that language learners cannot be proficient in a language unless they develop these three requirements. Firstly, it is vital to have language knowledge and be able to use its linguistic components appropriately. Then, it needs critical thinking and meta-cognitive skills in order to use the prerequisites, the acquired knowledge, and culture in specific contexts and communicative situations.

The American Council on the Teaching of Foreign Languages (ACTFL) Proficiency guideline presents proficiency as the language user's capacity to use language accurately and appropriately in authentic situations and natural interaction within a non-rehearse context (Lisa A. Harris et al. 2015).

For university EFL learners, English Language Proficiency (ELP) is the ability to communicate and interact effectively using Academic English in an authentic context. This ability is tested for various reasons including admission in a course of study, during, or at end of a course of study to test achievements. (TEQSS Agency, 2013). It can be developed through:

- Developing learners' academic literacy;
- Engaging learners in the course of study and the assessment process;
- Attaining all the learning objectives without any further support using other languages; and

- Using the English language effectively to interact in real-world situations outside the classroom.

The term “proficiency” in foreign language learning and teaching refers to *learners' accurate and appropriate use of a target language in both oral and written forms in different contexts and situations* (Cloud, Genesee, & Hamayan, 2000). It includes different levels that aims at classifying language learners into different categories according to their language skills' qualifications.

2.9.2 Proficiency Levels

Proficiency levels are a set of classifications that were created to categorise language learners based on their skills' performance. according to Oxfords dictionary (2010) Proficiency level is: A description of a language learners' level of performance in a target language, often described in terms of beginner level, intermediate level, or advanced level. Proficiency levels are often used to describe difficulty levels of language courses and language teaching materials (p.464).

Proficiency levels have been developed so as to distinguish between the different qualifications of the language users. Different frameworks have been developed to be used in foreign language teaching, learning, and assessment. However, the most influential frameworks were presented by the ACTFL and the Council of Europe (Maureen, 2009).

ACTFL (American Council on Teaching Foreign Language) presented a descriptive framework and a scale of proficiency that distinguishes between four levels of language users: novice, intermediate, advanced, and superior (Vossoughi and Javaherian, 2000). For EFL learners, developing these levels of proficiency does

not include reaching native speakers' fluency. However, being proficient includes the ability to use the language appropriately and accurately in regular classroom interactions and communication without help or support (Brown, 2001).

The Common European Framework of Reference (CEFR) for Foreign Languages is one of the most prominent frameworks of proficiency levels used for language teaching, learning, and assessment. It was produced and presented by the Council of Europe in 2001. This test is an international standard test that aims at labelling the language users' abilities. Unlike the ACTFL framework that distinguishes the levels in forms of descriptive labels, the CEFR presents six language proficiency levels in forms a range of letters from A1 to C2. It follows an ascending order (A1, A2, B1, B2, C1, C2) that starts up from A1 for the lowest level "beginners", to end up in C2 for proficient users "advanced" (Council of Europe, 2001). The levels are described in the following table:

Table 3
(CEFR 3.3): Common Reference levels - Global scale

Proficient User	C2 Proficient	Can understand with ease virtually everything heard or read, summarize information from different spoken and written sources, and very fluently and precisely express him/herself.
	C1 Advanced	Can understand a wide range of demanding, longer texts, and recognize implicit meaning, fluently and spontaneously express him/ herself, use language flexibly and effectively for different purposes and produce clear, well-structured, detailed text on complex subjects.
Independent User	B2 Upper Intermediate	Can understand the main ideas of complex text on both concrete and abstract topics, interact with a degree of fluency and spontaneity, and produce clear, detailed text on a wide range of subjects.
	B1 Intermediate	Can understand the main points of clear input on familiar matters, produce simple connected text on topics which are familiar or of personal interest, describe experiences and events, dreams, hopes & ambitions and briefly give reasons and explanations for opinions and plans.
Basic User	A2 Elementary	Can understand sentences and frequently used expressions, communicate in simple and routine tasks and describe in simple terms aspects of his/her background, immediate environment and matters in areas of immediate need.
	A1 Beginner	Can understand and use familiar everyday expressions and very basic phrases, introduce him/herself and others and can ask and answer questions about personal details and interact in a simple way provided the other person talks slowly and clearly and is prepared to help.

(Source: Council of Europe, 2022)

2.9.3 Assessing proficiency

The need to test and develop the language abilities and skills among foreign language users has led to the development of various Proficiency tests. According to Valette (1977), the main objective of a proficiency test “is to determine whether this language ability corresponds to specific language requirements” (p. 6).

Some of these tests have been developed and standardized for international use. Among the most prominent tests:

- The International English Language Testing System (IELTS)
- The American tests, the TOEFL, and the English Language Proficiency Test (ELPT)
- the Cambridge Certificate of Proficiency in English or CPE
- Test of English for International Communication TOEIC

2.10 Developing students' language proficiency through assessment

Developing learners' language skills towards proficiency requires language teachers to adopt several strategies in their instruction and assessment practice. Among these strategies, according to Prosser & Webb (1994, p.192):

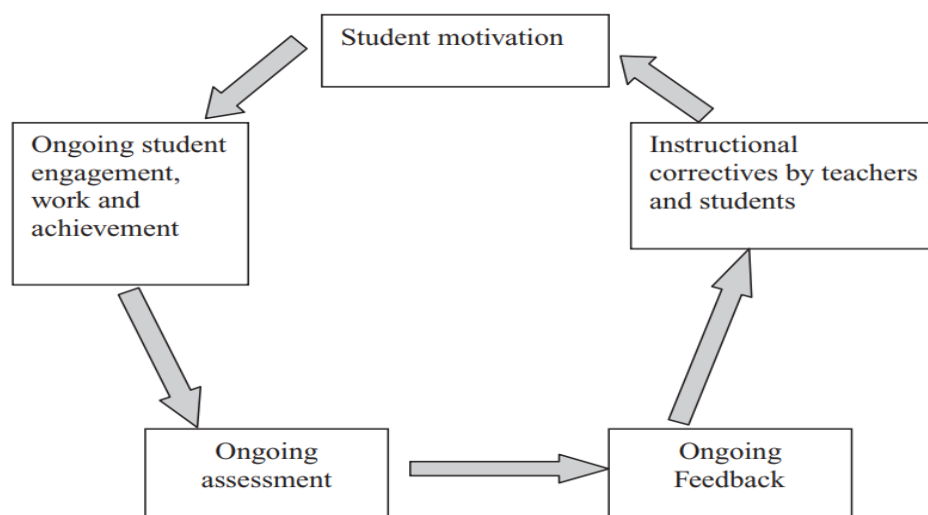
- Engaging the learners in active learning to create autonomous students who can actively learn and enhance their levels.
- Employing clear and comprehensible tasks.
- Selecting direct and clear learning objectives
- Sharing with learners the assessment expected criteria and rubrics which explain in detail how they will be evaluated and marked.

- Providing models and exemplars before the task which meet all the expected criteria.
- Using performance-based activities that help to develop the language and scaffold the learning process.
- Providing thorough and constructive feedback.

The use of assessment goes beyond evaluating and measuring learners' knowledge and skills. According to Wiggins (1998) tests can be educative in nature. Therefore, opting for the best assessment practices in EFL classrooms can play a vital role in supporting the learning processes and developing learners' skills (Stephenson, et al. 2003). Formative assessment, for instance, is an ongoing process of evaluating and providing feedback to students that can assist learners to enhance their performance and develop their proficiency levels. As illustrated in figure 3 below:

Figure 6

Formative assessment cycle.



Note. (a formative assessment cycle adapted from. Cauley & McMillan, 2010, p. 3)

According to Cauley & McMillan (2010), formative assessment can be used as a key element in enhancing learners' performance and skills, as well as fostering their degree of motivation. It provides valuable information about their performance throughout the courses. Based on this obtained information learners can recognize their flaws, reinforce their strengths and enhance their performance. Not just that, but the valuable information obtained from this type of assessment can also assist language teachers to refine their instruction and adapt what can help their learners to achieve better results. Accordingly, a remarkable improvement in learners' levels can be achieved over time. They confirmed that clearly stating that:

Formative assessment's emphasis on instructional modifications and student improvement supports student motivation and enables them to maintain high engagement and achievement. Using formative assessments effectively is indeed key to student motivation and achievement (P.6).

In the same regard, Bloom claimed that assessment can affect strongly both the learning and teaching processes, stating that:

Evaluation in relation to the process of learning and teaching can have strong positive effects on the actual learning of learners as well as on their motivation for learning and their self-concept in relation to school learning...Evaluation which is directly related to the teaching-learning process as it unfolds can have highly beneficial effects on the learning of students, the instructional process of teachers, and use of instructional materials by teachers and learners" (Bloom, 1969, p.50).

Additionally, effective feedback provided throughout the formative assessment process can be used as a powerful tool to enhance all learners' skills and performance, achieve better results and foster their motivation (Learning & Teaching Centre, 2013).

Conclusion

Based on all what have been mentioned in this chapter, we can say that the IPA (the integrated performance assessment) is a pedagogic and an assessment model that includes all the strategies and requirements needed to assess and enhance learners' language proficiency. Therefore, it was opted by the researcher in the present study as an alternative assessment model, firstly, to be used as a formative assessment for a semester long to test learners' language skills, secondly, to be implemented as a pedagogic and formative assessment tool to enhance their language proficiency.

The next chapter explains in details the research methodology opted for this study, presents the data collection instruments and procedures undertaken by the researcher to test the present study hypothesis and answer the study questions.

Chapter Three:

Research Methodology, Data Collection Instruments, and Procedures

Introduction

This chapter offers a brief description of the research methodology opted for the present study. It attempts to present the various steps undertaken to achieve the intended objectives of the investigation. First, it discusses some basic concepts about the research methodology framework that rationalises the choice of the methods of data collection and a description of the experimental design. Moreover, it provides a description of the experimental procedures, the selected population, the sample, and data collection instruments, as well as the statistical tools used to analyse and interpret the collected data.

3.1 Research Design and Methods

As stated earlier in the general introduction, this research aims to examine the effectiveness of using the IPA to assess and promote EFL students' language proficiency. It also seeks to examine the impact of such an assessment practice on students' motivation and attitudes. Thus, a number of methodological procedures were opted that will serve to effectively answer the research questions.

In the grounding pre-experimental phase of the study, the researcher wanted to understand the current assessment practices and teaching methods and strategies used in oral expression sessions. The qualitative research is a research method that is used with the aim of obtaining an in-depth comprehension of how different people behave, think, or feel towards a specific phenomenon (Kothari, 2004). It is thought it can be the best method to achieve this objective. Moreover, and because the researcher aims not only to explore current practices, but also to understand them and detect the possible flaws and try to address them through the suggestion of the IPA model, qualitative research can serve as a good starting point to investigate and

explore this phenomenon. This is due to the fact that descriptive research helps to establish a standing ground upon which the researcher can go deeper in research (Ahmed et al., 2019)

In the second step of the study, which attempted to test the study hypotheses, the experimental method was seen as the most appropriate method to use. This is because the experimental method serves to investigate the cause-and-effect relationship between variables (Creswell, 1994). Because this study sets a number of hypotheses, quantitative data measurement and analysis help to confirm or confute the study hypothesis by means of experimental or quasi-experimental designs (Creswell, 1994).

For this study, both approaches are crucial and combining them was necessary, therefore, a Mixed-methods approach was selected. This approach is also the appropriate choice to collect data so as to answer the research questions. According to Creswell and Plano (2011), A Mixed-methods approach is considered as a research method itself. It combines the qualitative and the quantitative methods under one research goal.

The rationale behind the use of such a method relies on the fact that it offers various supports to the research and the study findings. According to Dawadi, et al., (2021), Mixed-methods research has six major reasons for adopting it in a research study. The latter are summarised as follow:

- It integrates different types of data collection tools.
- It offers the researcher the opportunity to visualise the phenomenon under investigation from divergent perspectives and different angles.

- It increases reliability and validity of the findings.
- It can help the researcher to gain in-depth comprehension of the study inquiry and a deeper understanding of the whole study.
 - It provides a more holistic image of the treated phenomenon and complementary data to answer the study questions.
 - It can help in validating all conclusions and adding credibility to the study inferences.

3.1.1 The Experimental Design

This study followed an exploratory mixed methods sequential design of research. Within this framework, an experimental pre-test and post-test control group design was selected in order to test the effectiveness of the IPA in assessing and enhancing EFL students' language proficiency, in general, and their communicative skills, specifically, throughout three modes of communication: interpersonal, interpretive, and presentational. As it stands, the manipulation of the independent variable entails developing an IPA model, which serves as a classroom-based assessment, then test its effectiveness in improving EFL learners' language proficiency.

3.1.1.1 The Independent and Dependent Variables

Defining the experimental variables is an avoidable step in describing the selected experimental design used in this study. In experimental research, variables can be characteristics, condition or behaviours that can effect or be affected in quantity or quality (Creswell, 2012). During a treatment or an intervention of an experiment, researchers define and work with two variables. They

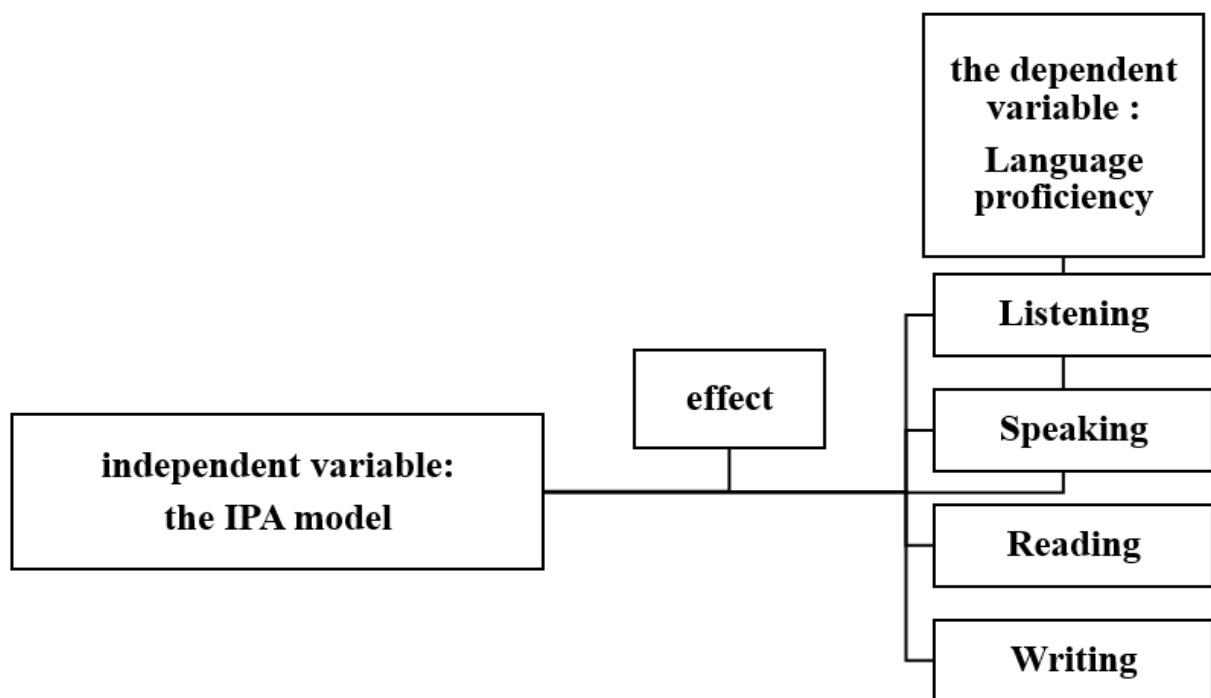
typically manipulate an independent variable, and measure a dependent variable so as to examine cause-and-effect relationships.

The independent variable is the cause or a treatment that should have effects on the dependent variable that is administered and manipulated by a researcher in order to discover, confirm, or refute its effects (Lammers & Badia, 2004). In the present study, the independent variable is the IPA model that involves three phases: the interpretive, interpersonal, and presentational modes of communication. The IPA model was implemented in a time span of one semester long (three months) with a sample of first year undergraduates with the use of a set of activities under each phase. The activities were designed and implemented in an attempt to cause improvement in students' performance in the oral expression session.

On the other hand, the dependent variable is the behaviour or the characteristic that receives the treatment or the manipulated effect of the independent variable in an experimental study (Lammers & Badia 2004). The researcher measures the dependent variable in order to explore these effects. Within the present study's scope, the dependent variable is the EFL students' language proficiency, in general, and, specifically their oral and spoken skills. As it has been stated in the second chapter, language proficiency is represented through the four language skills: listening, speaking, reading and writing. Figure 7 explains this study's variables

Figure 7

The relationship between dependent and independent variables



3.1.1.2 The Pre- and Post-test Control Group Design

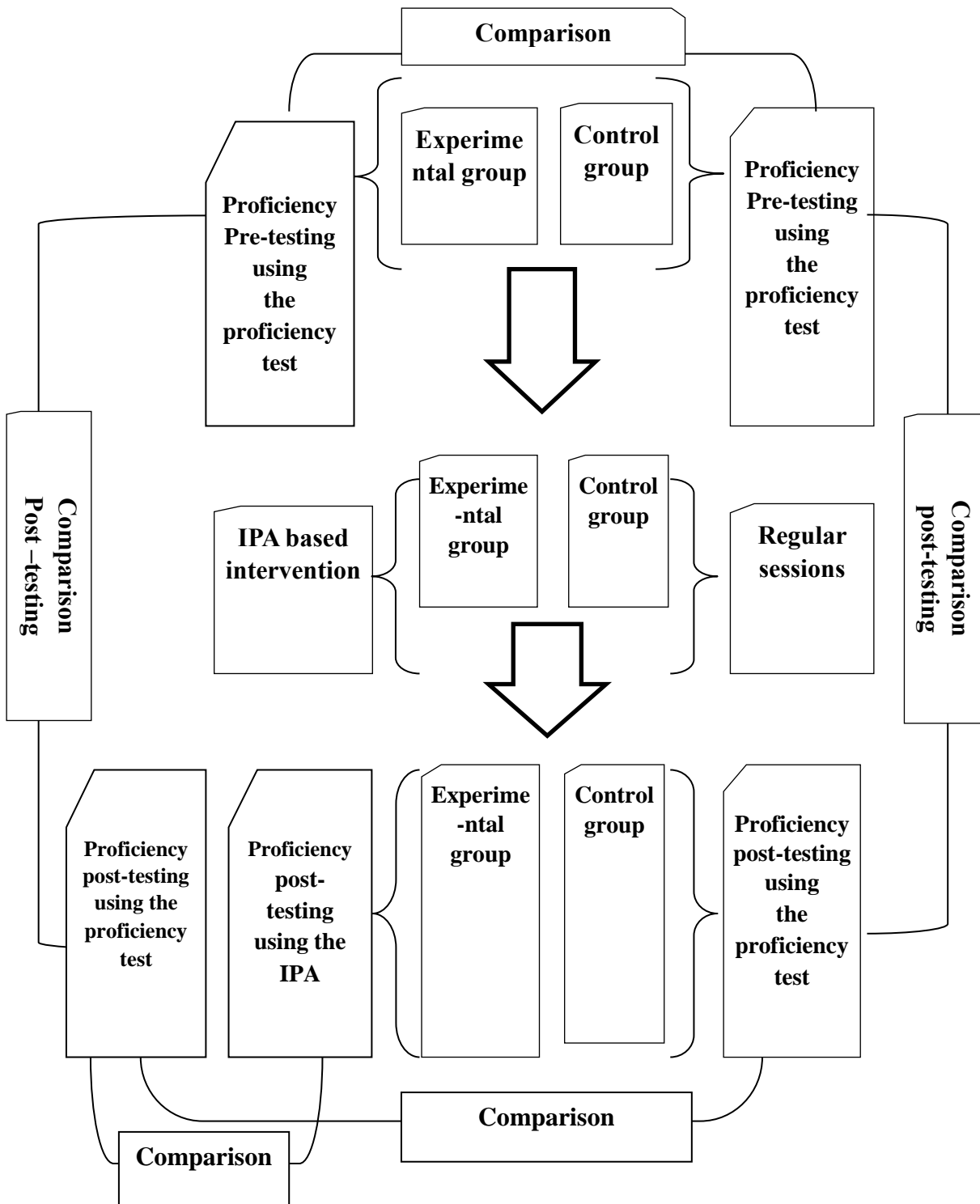
The quasi-experimental design is a type of experimental research that makes use of an ‘intervention’ in forms of a treatment within a specific programme or a policy in order to test to what extent it can attain its intended purposes (Rogers et al., 2019). It resembles the true experiment, however, a quasi-experimental design lacks at least one of its defining conditions.

Quasi-experimental design is used for research when it is difficult or impossible to assign participants through random selection (Martella, 2013, p. 150). Yet, assignment to conditions can be either by self-selection where the participants choose to be involved in the experiment or by administration where the participants are already selected by the policy maker, teacher, officials or so on. Otherwise, both of these methods can be used (White & Sabarwal, 2014).

Accordingly, and since this study used non-random convenient sampling, it lacks one of the true experimental identifying conditions. However, the researcher assured the existence of the other conditions, including the homogeneity of distribution in both groups, as well as normal distribution of the proficiency levels.

Since this design is a two groups design, a number of comparisons were conducted before and after the implementation of the treatment with the experimental group. The first step in the experimentation is the pre-experimental phase in which control and experimental groups were pre-tested to confirm the compatibility of levels and that both groups share the same baseline traits. In the second stage the two groups received oral expression sessions. The experimental group received the IPA intervention while the control group received regular oral expression sessions. At the pre-intervention a post-testing is conducted among the two groups to identify the impact of the IPA intervention. Accordingly, comparisons between both groups pre and post performances were run. Figure 8 represents the experimental procedures conducted in this study's research design.

Figure 8
Experimental procedures



3.2 Research Setting and Participants

This study was carried out in the English Language Department at M'sila University. It was conducted during the second semester of the academic year

2021/2022. It lasted for three months (from March 08th, 2022- to May 24th, 2022), including nearly 11 sessions. These sessions were held in the EFL Department's Laboratories where teachers are provided with all the materials needed for the oral expression sessions.

The population of this study are EFL Algerian university students. Yet, the accessible population were LMD students at M'sila University. The selected sample was conveniently formed from first year LMD undergraduate students. Thus, an inescapable non-random sampling method was used due to the circumstance in which the study was carried out. The latter include mainly the health protocol, and the resulting teaching and scheduling routines. Accordingly, the researcher was obliged to work with the groups she was assigned to teach.

Consequently, the study sample included 52 first year undergraduate EFL students divided into an experimental and a control group. In addition, the participants were EFL students majoring in English language. They have been taught English language for seven years before being qualified to admission at the English language Department. This means that the participants of this study are estimated to have a good grasp of the English language, which would assist them to understand all instructions of the IPA tasks, as well as the theme and the content.

As to the teachers' sample participating in the grounding phase of the study, 35 university teachers responded to the online questionnaire. Thus, the teachers' sample was haphazardly designated through responding to the online questionnaire that was directed to Algerian EFL university teachers, specifically, oral expression teachers. Since the questionnaire was online, teachers were from different Algerian

universities. Table 4 involves a summarised description of the participants in this study.

Table 4.

Study Participants Description

Sample	Information	Count	%	
Teachers	Gender	Female	18	19,51
		Male	17	20,73
	Rank	Master	10	12,20
		Magister	12	14,63
		Doctorate	13	15,85
		Teaching experience	1-5 years	12
5-10 years			15	18,29
More than 10 years			8	9,76
Piloting sample		Raters	2	2,44
		Experts	5	6,10
Students		Group	Experimental group	26
	Control group		26	31,71
	Piloting sample		25	30,59

In addition to the main study sample, a piloting sample was also needed. It involved five teachers, who served as experts to evaluate content validity of the questionnaires, two teachers served as raters in examining inter and intra rater reliability, and thirty students served to pilot the IPA model and the attitudes questionnaire validity and reliability.

3.3 Data Collection Instruments

The instruments used to collect data in this study included two questionnaires, an IPA based intervention, and a Proficiency Test. The first questionnaire was designed and devoted to investigate current perceptions of assessment practices among EFL University oral expression teachers. The second questionnaire was dedicated to explore students' attitudes towards the IPA model use. The proficiency test was used as pre- and post-test students' overall language proficiency.

3.3.1 Teachers Perceptions of and Practices Questionnaire

As already explained in prior sections, the aim behind the use of teachers' questionnaire is to explore teachers' perceptions of and practices of teaching and assessment in oral expression module since it was selected to be the experimental context. It was also designed to test their perceptions about the integration of the four skills in both teaching and assessment in oral expression module and their readiness to use the IPA model. Therefore, the questionnaire involved many types of items both open-ended and closed ended. The items were distributed into three main sections. Table 5 represent teachers' questionnaire sections.

Table 5.

Teachers' questionnaire sections and items

Sections	Title	Number of items
Section1	General information	3 items
Section2	Teachers' perceptions and practices towards oral expression teaching	7 items
Section3	Perceptions and practices towards oral expression Assessment	10 items

The first section items were set to collect information needed to describe the study sample. The data obtained from the first section are presented in Table 6, in the sample's description section of the Methodology Chapter. The second section items were designed to probe for teachers' current practices regarding oral expression teaching. The first item (item4) in this section explored the teachers' perceptions about the objectives of teaching oral expression. The second item (item5) sought to understand the language skills that are the focus of oral expression module. Then, items 6 and 7 sought to shed light on the most used activities in the oral expression session and the most enjoyed ones. Item 8 was dedicated to investigate whether students were involved in selecting the contents and activities of the oral expression session or not.

As for the investigation of the difficulties teachers face when teaching oral expression, item 9 is set to do this purpose. The last item (item 10) targeted one of the core concerns of this study which is the integration of the four skills in the oral expression session and the methods through which teachers use the four skills in their sessions.

The second section main objective was to understand teachers' perceptions of and practices vis-à-vis the assessment in the oral expression module. It started by asking teachers about the assessment methods applied in the oral expression session (item11). Items 12, 13 and 14 focused on investigating the feedback practices among oral expression teachers through asking them about the frequency of providing feedback to students (item12), the type of feedback provided (item 13), and the method of providing this feedback (item14). The types of mistakes targeted by teachers' feedback were the core concern of item 15. Moreover, to understand the

possible difficulties teachers face when assessing students' performance in an oral expression session, item 16 was designed. As for the integration of the four skills in the assessment process, items 17 and 18 were designed to ask teachers about their use of the four skills and their perceptions about its effectiveness. The last two items (item 19 and item 20) were devoted to examine teachers' perceptions about alternative assessment importance (item 19), and their use of these methods in the oral expression sessions, in addition to their knowledge about these methods.

3.3.1.1 Teachers' Questionnaire Reliability and Validity

To check the validity of the teacher questionnaire, a pilot study was carried out before administering the questionnaire with the main study sample. The targeted types of validity were face and content validity. Thus, the first draft of the questionnaire was piloted with five teachers. The piloting sample was invited to examine each item and specify whether it is relevant or not to the objective of the tool as part of the content validity analysis. Face validity was achieved since all the five experts agreed on the general relevance of the sections and items types and number to the intended objectives. The collected data from the experts' opinions provided valuable information that helped in revising and developing the final version of the questionnaire. To examine content validity, an evaluation sheet was attached to the questionnaire to collect necessary feedback. The piloting sample feedback and comments on the questionnaire two sections are summarised in table 6 and table 7.

Table 6 summarises the experts' agreements and disagreements about section 1 items. Items that received a percentage of agreement of 80% or 100% were

accepted as being valid. Items that received 60% were modified while items that received 40% and 20% were removed being irrelevant.

Table 6.

Content Validity of Teachers' Questionnaire Section 1

Section	Item	Agreement	Disagreement	Remarks
1st	4	100%	0%	Valid
	5	80%	20%	Valid
	6	60%	40%	To modify
	7	80%	20%	Valid
	8	80%	20%	Valid
	9	60%	40%	To modify
	10	0%	100%	To remove
	11	100%	0%	Valid
	12	20%	80%	To remove

From table 6, it is indicated that the first version of the questionnaire's first section included nine items. Apparently, most of the items received 80% to 100% of agreement and thus are valid for the investigation. Item 6 and 9 received 60% of agreements thus were set to modification, as suggested by the experts. Item 6, for instance, was as follows "What is the teaching method you follow when teaching oral expression?" it was replaced by "What are the most used activities in your oral expression session?" being more practical and clear and since most teachers are not aware of the teaching methods they may be using and because the teaching activities are more common among teachers of different years of experience. As for item 9, it was a close ended item that the experts suggested that it is better to replace it by an open ended item. Two items were removed as, according to experts' disagreement, item 10 received 100% disagreement and item 12 received 80% of disagreement.

Similarly, section two items validity was examined in terms of experts' agreement. Table 10 represents section two items.

Table 7.

Content Validity of Teachers' Questionnaire Section 2

Section	Item	Agreement	Disagreement	Remarks
2 nd	13	100%	0%	Valid
	14	0%	100%	To remove
	15	100%	0%	Valid
	16	80%	20%	Valid
	17	80%	20%	Valid
	18	100%	0%	Valid
	19	100%	0%	Valid
	20	60%	40%	To modify
	21	80%	20%	Valid
	22	60%	40%	To modify
	23	100%	0%	Valid

According to table 7, the first version of the second section included 11 items. Among the 11 items, item 14 was removed as it received 100% disagreement confirming its irrelevance to the questionnaire contents. Items 20 and 22 were modified, building on experts' opinions. Item 20 (in the final version is item 17) was changed from "How often do you use the four skills in assessment?" to "Do you integrate the four skills when evaluating oral expression?". Item 22 (in the final version is item 19) was modified from "Why are traditional method of assessment less effective than alternative ones?" to "Do you think that traditional types of assessment are sufficient to test students' language proficiency?" because the first

version of item 22 seems judgmental and guides teachers' responses to a specific answer.

3.3.2 Students Questionnaire

After the end of the experimental phase, the experimental group were asked to respond to the attitudes questionnaire. It attempted to test the washback of the IPA model on their attitudes towards the oral expression module and the assessment model used in the second semester. The implemented questionnaire was a structured questionnaire in which all items were close-ended Likert scale items. The total scale involved 19 items. Table 8 represents the questionnaire's final version structure.

Table 8.

Students' Attitudes Questionnaire

Section	Title	Number of items
1 st section	Students' attitudes towards oral expression module	11 items
2 nd section	Students' attitudes towards IPA activities and outcomes	8 items
Total		19 items

The Final version of the students' questionnaire comprised two main sections. The first section was devoted to probe EFL students' attitudes towards the different areas of the oral expression teaching, including attitudes towards the contents and activities (item 8, item 7, item 1, item 3, item 11), assessment (item 2), the use of the reading and writing skills (item 6), motivation during the sessions (item 5) teachers' feedback (item 9, item 10), as well as their attitudes towards the first semester oral expression session (item4). The second section tested the students' attitudes towards

the aspects related to the implemented IPA model. It explored the students' attitudes about the activities and phases of the IPA (item 12, item 13, item 17), the skills integration (item 14, item 18), feedback provision (item19, item16, item15).

3.1.1 Students Questionnaire Validity and Reliability

The first version of the questionnaire was piloted to examine its face and content validity. Similar to teachers' questionnaire, the students' questionnaire was examined by the five experts. It was not possible to conduct reliability analysis, namely, Cronbachs' Alfa and internal consistency because of the nature of the attitudes questionnaire that was targeting students who received an IPA intervention. Experts' agreement levels on the first version of items are summarised in table 9.

Table 2.

Experts Agreement on Students' Questionnaire

Section1			Section2		
Item	Agreement %	Decision	Item	Agreement %	Decision
1	80	Valid	12	100	Valid
2	100	Valid	13	0	To remove
3	100	Valid	14	80	Valid
4	40	To modify	15	80	Valid
5	100	Valid	16	40	To modify
6	80	Valid	17	100	Valid
7	60	To modify	18	100	Valid
8	60	To modify	19	20	To remove
9	80	Valid	20	60	To modify
10	100	Valid	21	0	To remove
11	100	Valid	22	80	Valid

The first version of the questionnaire included 22 items. Three items were removed (item 13, item 19 and item 21) since they received low agreement from the experts. Three items from section 1 were modified and two items from section 2. The final version, then, included 19 items.

3.3.3 The IPA Model Based Intervention

The basic aim behind an experimental research relies on testing the effectiveness of a given stimuli on a certain variable, in a controlled environment that prevents the influence of any other possible interfering factors in order to get valid results (Creswell, 2003, p. 154). Accordingly, the treatment or the intervention of an experiment involves a set of conditions under which the experimental and control groups are treated (Kothari, 2012, p.35). In the current study, the treatment was an innovative type of assessment model, that is "The Integrated Performance Assessment - IPA", designed so as to be applied with the experimental group.

As it has been mentioned so far, the treatment lasted for a semester long on which six sessions were dedicated to the implementation of the IPA in addition to two further warm up sessions. Each session lasted for about two hours. During these sessions, the students were integrated in three modes of communication using various tasks in the oral expression module to assess and boost EFL students' language proficiency.

The researcher chose to implement this innovative assessment model in the oral expression module because of two reasons. Firstly, IPA is designed to assess students' language skills, communicative skills, and proficiency, including (language fluency, accuracy, appropriateness, communication strategies, interaction, language form...), not the content of the modules. Therefore, the oral and the

writing expression modules could be the best choice since they are the only modules that have no specific content to deliver and they only test language skills. Secondly, the oral expression module includes “free topics” in its curriculum. i.e., the teacher is allowed to select any topic that suits the students' interest. This was appropriate for the IPA because this assessment model also should be involved in a theme of the students' interest. For the planning of the implementation, the treatment timing was scheduled as illustrated in table10.

Table 3.

IPA based intervention schedule and allotted times

Theme	Number of sessions	Allotted time	Date
Tourism	7session + 2 additional sessions	14 hours + 4 hours	From: March 8 th To: May 24 th , 2022.

The content materials including the videos and the texts were carefully designed and selected from authentic resources created by native speakers. The selected themes were consistent with the provided syllabus in the English language department. Since the second semester syllabus was devoted to free topics, the researcher selected a topic of interest to students after asking them about their preferences and needs.

The IPA model was divided into three phases, “Interpretive, interpersonal and presentational modes of communication”. Each phase was carried out in two sessions, one for the tasks of the modes of communication, and the other one for feedback and correction. To ensure a full understanding of the tasks of the IPA, the researcher provided a general introduction and an overview that described for the student the context and purpose of the series of the authentic tasks. The researcher

also provided the students with a detailed explanation of the assessment model, clarified the link between the tasks of all the three phases and explained how each task is the standing as a ground for the subsequent in order to accomplish the final oral presentational product that would be their summative assessment for the module. Table 11 represents the information related to IPA model structure.

Table 4.
IPA model structure and objectives

Phase	Teaching activities	Feedback type	Number of sessions	Objectives
<i>Interpretive</i>	Reading a text then doing Reading Comprehension tasks	Feedback on Interpretive Task (Review of rubrics; Discussion of errors)	2 sessions	<ul style="list-style-type: none"> ✓ <i>Help students develop their knowledge repertoire of the selected theme and</i> ✓ <i>Enhance their Reading Skill</i>
<i>Interpersonal</i>	<ul style="list-style-type: none"> • Task 1: Watching a short video to refresh students' minds about the topic doing a listening comprehension task. • Task 2: Interacting orally about the theme one by one and answering questions about what have been watched 	Feedback on both tasks of Interpersonal phase (Review of rubrics; Discussion of errors)	2 sessions	<ul style="list-style-type: none"> ✓ <i>Use the adopted information and knowledge to interact orally and participate in classroom discussions</i> ✓ <i>Enhance their listening and speaking Skills.</i> ✓ <i>Spot the light on students weaknesses and language defects.</i>
<i>Presentational</i>	From all what have been learned in previous sessions, Prepare a presentation about the theme (using audio visuals and oral presentation) + And a written copy of the work.	Feedback on Presentational Task (Review of rubrics; Discussion of errors) of both the presentation and the written work.	2 sessions	<ul style="list-style-type: none"> ✓ <i>communicate idea and personal views to audience through oral and written products</i> ✓ <i>enhance their speaking and writing skills</i>

The researcher prepared lesson plans for each session specifying the short term objectives, allotted time, integrated skills, materials used and feedback. First, the

researcher started each phase with a warming up that was done in a whole session in which she prepared the students and stimulated their interests to do the targeted tasks. The warm up of the first phase was an introduction to the IPA where every single detail of this assessment process was explained and at the same time, the students helped to select the theme and the topic. The second warm up session was conducted after the spring vacation to refresh again the students' minds and integrate them again to the IPA process. The final warm up session was after the holiday of Aid El Fitr. Additionally, the researcher tried to create a supportive learning environment so as to integrate the students in the mode of the communication in each phase. The following sections provide an explanation of the three phases of the model and related activities.

3.3.3.1 The Interpretive Phase

The interpretive mode, as the first phase on the IPA, implicates literal comprehension and interpretation, as well as making inferences "reading between the lines". Accordingly, the students are required to identify cultural perspectives and participate in classroom discussions to express their personal opinions (Shrum & Glisan, 2010). During this stage, students read a text about "Tourism in London" that helped these students to gain critical knowledge and familiarity of the content and the theme treated. After the reading phase, the students were engaged in class discussion where the teacher was only guiding the discussion and helping to generate the ideas.

Table 12 taken from the IPA model represents the Integrated Performance Assessment Comprehension Guide Analysis: Performance Expectations.

Table 5.

Integrated Performance Assessment Comprehension Guide Analysis: Performance Expectations.

Criteria	What is involved
Literal Comprehension	
Key words recognition	Students identify key words in the text that provide clues to the overall meaning of the text.
Main ideas detection	Students identify the main ideas of the text.
Supporting detail detection	Students identify important details that further explain the main idea(s).
Interpretive Comprehension	
Organizational features	Students identify the ways in which the text is organized and the purpose of organizing the information in that manner.
Guessing meaning from context	Students use contextual clues and knowledge of language and text structure to infer the meaning of new words and phrases.
Inferences (Reading between the lines)	Students interpret the overall meanings contained in the text by combining knowledge of key vocabulary, important details, text features, and their own background knowledge.
Author's perspective	Students identify the author's perspective and provide justification.
Cultural perspectives	Students identify cultural perspectives/norms by connecting practices/products to cultural perspectives.

At the end of this stage, the students were provided with worksheets entailing reading comprehension tasks about the text to answer then deliver to the teacher for feedback and correction. The worksheets were designed based on the IPA Comprehension Guide Analysis extracted from the IPA manual. The reading

comprehension questions included key word detection, main idea and supporting detail detection, drawing conclusions, and inferences, identifying organisational features of the text, guessing meaning from context, making inferences, and discovering the author/cultural perspectives described in the table above.

The second session of the interpretive phase was dedicated to Responsive, and Co-constructive Feedback that aimed to improve student performance. During this session, the students received their corrected worksheets and negotiated the text's inferences and the difficult tasks they faced with the teacher.

The feedback was a significant opportunity to help these students realise their weaknesses whether in interpreting, problem-solving, or in reading strategies. The teacher tried to help the students to solve the difficult tasks by using assessing and assisting questions to allow them to participate and realise their flaws. At the end of the session, each student was provided with a chance to evaluate the accuracy of their own and others' contextual guessing in a class discussion with the assistance of the teacher.

The scoring rubric was divided into two parts. Part one was for literal comprehension. It was scored on nine points, whereas, the interpretive comprehension part was scored on eleven points.

3.3.3.2 The interpersonal Phase

The second phase of the treatment included an active negotiation of meaning and personal views on the same theme. The first session was a warm up session, conducted after the spring vacation, to refresh the students' minds about what has been seen in the previous sessions and prepare them to the interpersonal phase. The

teacher started the session with a video watching about "Tourism in London", which was an authentic created by native English speakers. After watching the video, the students were required to answer listening comprehension task in worksheets.

The second session, which represents the main interpersonal phase, was devoted to a speaking task. The teacher started open discussion with the students on the same theme, and then asked about the students' their opinions about "tourism in London", tourism in Algeria, and the best places for tourism according to them. During the discussion, all the students were involved and asked about their views. The teacher guided the discussion by asking questions, by which she evaluated the students' language, speaking performance, and comprehension. Since the topic was of their interest, remarkable interaction and participation were noticed. To perform this task, the students communicated and interacted with each other spontaneously with no use of any written support. Their oral responses and answers were evaluated and graded based on the table below "Interpersonal Mode Rubric" provided by the IPA manual. The total score obtained from the five evaluated aspects of the language **"Language Function, Text Type, Communication Strategies, Comprehensibility and Language Control"** is twenty (20) points.

Table 6.*Integrated Performance Assessment (IPA) Interpersonal Mode Rubric*

Language Function	<i>A language task on which the speaker is able to handle in a consistent, comfortable, sustained, and spontaneous manner</i>
Text Type	Quantity and organization of language discourse (continuum: word - phrase - sentence - connected sentences – paragraph - extended discourse)
Communication Strategies	<i>Quality of engagement and inter-activity; how one participates in the conversation and advances it; strategies for negotiating meaning in the face of breakdown of communication</i>
Comprehensibility	Who can understand this person's language? Can this person be understood only by sympathetic listeners used to interacting with non-natives? Can a native speaker unaccustomed to non-native speech understand this speaker?
Language Control	Grammatical accuracy, appropriate vocabulary, degree of fluency

Source: (Adair-hauck et al., 2016)

The teacher provided another portion of co-constructive and descriptive feedback concerning the students' performance and language instantly after each presentation. She tried to give a chance to all the students to assess their performance individually with her by assessing and assisting questions to co-construct effective feedback. This interpersonal task took place in two sessions because of large number of the students within the group.

At the end of this phase, the teacher presented an integrated situation to the students and asked to prepare a presentation based on the two previous phases in which the students used all the adopted repertoire of the linguistic skills and knowledge to be presented in the next phase. The students were asked to imagine themselves being tourist guides working in London or any touristic city of their choice, promoting a tour there using information about the city's history, famous landmarks, and attractive qualities. Correspondingly, they had to use the information they had obtained from all the previous tasks to convince tourists to visit London.

The presentations were required to include the following points accompanied by short videos, PowerPoint presentations PPT, or pictures related to the theme:

1. A formal greeting;
2. introduction to the tourists about the touristic plan;
3. introducing the target landmark's name and location;
4. describing the place around;
5. providing a historical overview about each of the selected landmarks;
6. suggesting famous restaurant where they can find famous food a memorable meal, how much it cost and your opinion of it. (imagine)
7. how you can get around in the city. (Transportation) ; and
8. your overall opinion about the city its Landmarks.

3.3.3.3 The Presentational Phase

The third and last phase of the IPA model was the presentational phase. This phase required the student to prepare a one-way presentation. During this session,

students were asked to present their final works in the form of oral presentations accompanied by PPT presentations or a video as it was mentioned earlier.

At the beginning of the session, a warm up activity was done to prepare the students again and involve them in the IPA after Aid El Fit vacation that lasted for 15 days. The teacher presented a modelling video to portray the expected performance because “Models set the standards that we want to achieve... they anchor the feedback” (Wiggins, 1998, p. 64). Additionally, she reviewed with them the basic evaluation criteria of this phase which were similar to the interpersonal phase criteria, except for one different aspect (Impact): “Language Function, Text Type, **Impact**, Comprehensibility, and Language Control”, as well as “fluency accuracy and form”

Table 7.

Integrated Performance Assessment (IPA) Interpersonal Mode Rubric

Language Function	<i>A language task on which the speaker is able to handle in a consistent, comfortable, sustained, and spontaneous manner</i>
Text Type	Quantity and organization of language discourse (continuum: word - phrase - sentence - connected sentences – paragraph - extended discourse)
Impact	<i>The clarity, organization, and depth of the presentation and the degree to which the presentation maintains the attention and interest of the audience “powerful, memorable, provocative, or moving”.</i>
Comprehensibility	Who can understand this person's language? Can this person be understood only by sympathetic listeners used to interacting with non-natives? Can a native speaker unaccustomed to non-native speech understand this speaker?
Language Control	Grammatical accuracy, appropriate vocabulary, degree of fluency

Source: (Adair-hauck et al., 2016)

While presenting their products, the students were allowed to use notes from time to time. After each presentation, the students were evaluated and graded based on the previously mentioned criteria in the above table provided by the IPA manual. The total score obtained from the five evaluated aspects of the language “**Language Function, Text Type, Impact, Comprehensibility, and Language Control**” was 20 points.

Additionally, the teacher offered feedback on students' performance in the form of questions so that the students construct their own feedback and recognise their own flaws and strengths. Also, the teacher asked the other students to give their feedback or ask the presenter questions on the topic.

3.3.4 The Language Proficiency Test

The researcher employed the IELTS academic test (of USA), in the present study, to test the students' English language proficiency variation after a semester long of using the IPA.

3.3.4.1 The IELTS Proficiency Test

The International English Language Testing System (IELTS) is one of the pioneering proficiency tests used in English-speaking countries, including: the USA and the UK. It aims to test foreigners' and migrants' English language proficiency for many reasons, such as admission to educational institutions, professional registrations, and approval or for immigrants. This proficiency test involves assessing the four language skills: Listening, Speaking, Reading, and Writing. The tests could be taken all at the same time with no breaks. Only the speaking test could be delayed or taken online (IELTS Website).

The rationale behind selecting the IELTS academic test, as a tool to measure the English language proficiency in this study, was for many reasons. Firstly, most of the previous studies that aimed to test the development of language proficiency and assess language skills employed the IELTS Academic tests as a testing tool (Archibald, 2001; Read & Hayes, 2003; Green, 2005; Craven, 2012). Secondly, it was selected because of its accessibility and availability in that we can find reliable samples with their correction on the IELTS official website. Additionally, since the IELTS is an academic test widely recommended and used by various international universities to test English Proficiency, it is likely to be best for this study. (Craven, 2012; Dooney, 1999; Feast, 2002; O'Loughlin & Arkoudis, 2009).

Finally, the IELTS tests are designed by specialists in a way that all these tests are fair, accurate, reliable, and valid for all test takers, in all countries and at all times. This ensures consistent and valid results (IELTS website, 2022). Thus, selecting tests from the available samples can be used in the present study with no need to for further confirmation of their reliability or validity.

The only limitation that could face the researcher when using the IELTS lies in the fact that the students with band score 7 (C1) at pre-test, by post-test may not increase to higher levels and they could even show lower scores. However, the students with lower scores will show remarkable improvement at the end of the post-test (Green, 2005).

Various previous studies, on which the IELTS was used on EFL classrooms to test students' proficiency or skills, used the adaptation and not the full version (Archibald, 2001; Green, 2005; Humphreys, et al., 2012). Hence, the researcher selected the tests from the samples provided by the official website of the IELTS

then adapted the original versions to what suits the level of the EFL first year students, as well as the time dedicated to the test.

3.3.4.1.1 Assessing the Listening skill

The listening test for the IELTS includes 40 questions divided into four sections with 10 questions in each. Thirty minutes are dedicated to listen then answer the questions. The answers follow the same order as of the listening scripts, which they hear only once. These recordings include different English accents, such as the American and the British. Additionally, each of the sections involves social contexts in the form of conversations and monologues. The following table summarises the listening test parts, questions, requirements and scoring.

Table 8.

Description of the IELTS Listening test.

Timing	Questions	Test Parts	Skills criteria	Marking
30 minutes to listen to four recordings and then answer the questions	<ul style="list-style-type: none"> • 40 questions. • multiple choice, matching, form completion, note completion, table completion, flow-chart completion, summary completion, sentence completion, short-answer questions. 	<ul style="list-style-type: none"> • 4 sections, in forms of dialogues, conversations, monologues... • In social, educational, academic or training contexts created by native speakers 	<ul style="list-style-type: none"> • understanding of main ideas • understanding of specific factual information • recognising opinions, attitudes and purpose of a speaker • following the development of an argument. 	<ul style="list-style-type: none"> • Each correct answer receives 1 mark. • The Scores obtained are transformed into the IELTS 9-band score.

3.3.4.1.2 Assessing the Reading skill

The reading tests is taken in 60 minutes. Test takers were required to read and answer 40 questions. Tasks consisted of different types of questions that included both literal reading comprehension and inferencing (Reading for gist, main ideas detection, finding detail, logical arguments' comprehension, and identifying writers' opinions and intentions...).

The reading tasks entail three texts Adopted from different resources, such as: books, articles, journals newspapers, and magazines. Types of these texts vary in each part of the reading test. They include descriptive, factual, discursive, and analytical.

Table 9.

Description of the IELTS Reading test.

Timing	Questions	Test Parts	Skills criteria	Marking
60 minutes Reading and answering questions	<ul style="list-style-type: none"> • There are 40 questions. • multiple choice, identifying information (True/False/Not Given), identifying a writer's views/claims (Yes/No/Not Given), matching, sentence completion, summary, note completion, table completion, short-answer questions... 	<ul style="list-style-type: none"> • 3 sections. • Using authentic texts taken from books, journals, magazines and newspapers. • topics of general interest. • Texts range from the descriptive and factual to the discursive and analytical. 	<ul style="list-style-type: none"> • reading for gist • reading for main ideas • reading for detail • understanding inferences and implied meaning • recognising writer's opinions, attitudes and purpose • following the development of an argument. 	<ul style="list-style-type: none"> • Each correct answer receives 1 mark. • Scores out of 40 are converted to the IELTS 9-band scale.

3.3.4.1.3 Assessing the Writing skill

The writing skill in the IELTS test is assessed based on four criteria (Task Achievement points + Coherence and Cohesion points + Lexical Resource points + Grammatical Range and Accuracy points). The test is conducted in 60 minutes through which test takers are required to accomplish two tasks. The tasks consist of writing two paragraphs on two different themes including 150 words for the first task and 250 words for the second task.

Table 10.

Description of the IELTS writing test.

Timing	Questions	Test Parts	Skills criteria	Marking
60 minutes to complete two tasks and covers topics of general interest	<ul style="list-style-type: none"> • 2 tasks. • 150 words for Task 1 • 250 words for Task 2 • Written in an academic, semi-formal/neutral style. 	<ul style="list-style-type: none"> • There are 2 parts. • Part 1: summarise or explain the information in your own words. • Part 2 write an essay in response to a point of view, argument or problem. 	<ul style="list-style-type: none"> • Task Achievement • Coherence and Cohesion • Lexical Resource • Grammatical Range and Accuracy. 	<p>the total score is calculated by an arithmetic mean:</p> <p>(Task Achievement points + Coherence and Cohesion points + Lexical Resource points + Grammatical Range and Accuracy points) /4 = final score.</p>

3.3.4.1.4 Assessing the Speaking Skill

The Speaking test in the IELTS is made of two parts in the form of face-to-face interview and interaction between the test taker and an examiner. It aims at testing the ability to communicate in English language appropriately. The examiners evaluate the test takers performance based on four 4 criteria: Fluency and coherence, lexical resource, grammatical range and accuracy, and pronunciation.

Part 1: Time takes between 4 to 5 minutes and includes general questions about the test takers such as questions about their families, countries, occupations, educational level and interests.

Part 2: It takes from 3 to 4 minutes during which the examiners present some topics to the test takers, they give them 2 minutes to think and then respond. They may also ask questions to help the test takers to finish the task.

Part 3: This part takes from 4 to 5 minutes. It follows the previous (Part 2) by giving further questions about the same topic. It aims to provide more opportunities to test takers to communicate their ideas and interact using the language.

Table 11.*Description of the IELTS Speaking test.*

Timing (14mins)	Test Parts (2 parts)	Questions (interaction and discussion)	Skills criteria (4 criteria)	Marking (arithmetic mean)
(4-5 mins)	Part one : Introduction	General questions on familiar topics, e.g. home, family, studies, work, interests, future plans, past experiences	<ul style="list-style-type: none"> • Fluency and Coherence • Lexical Resource • Grammatical Range and Accuracy • Pronunciation 	<p>the total score is calculated by an arithmetic mean:</p> <p>(Task Achievement points + Coherence and Cohesion points + Lexical Resource points + Grammatical Range and Accuracy points) /4 = final score.</p>
(3-4 mins)	Part Two : Long-turn	Talk about a given topic for 1-2 minutes. Prompts are provided on a task card and 1 minute preparation time is given for you to make notes and think about what you want to say.		
(4-5 mins)	Part Three : Discussion	Discussion type questions related to the topic of Part 2 where you are expected to express your opinion about more abstract issues and ideas.		

3.3.4.2 The IELTS in The pre-test and post-test

The instruments of the proficiency test were adapted from samples of IELTS test obtained from IELTS official website. It was administered as the first and the final step in the present research and used as a pre-test and post-test and adapted to what suits the first year students' levels. Therefore, both proficiency tests used in this study consisted of 74 different types of questions, including multiple-choice items testing vocabulary, structure, listening, and reading comprehension. The tests were divided into four phases to test the four skills.

The first phase took 30 minutes and was dedicated to testing the listening skill. It included 20 questions composed of multiple choice, matching, completion, note completion, and short-answer questions.

The second phase took 60 minutes dedicated to testing the reading skill. It was composed of three reading comprehension parts including three short texts with 40 questions (multiple choice, identifying information, identifying main ideas, sentence completion, and finding synonyms).

The third phase took 14 minutes and was dedicated to testing the speaking skill. It was composed of 12 questions divided into three parts: Introduction, Long-turn, and Discussion.

The fourth phase took 60 minutes and it was dedicated to testing the reading skill. It was composed of two writing parts: part one includes reading and summarizing a text while part two was devoted to writing an essay.

3.4 Procedures

In light of what has been mentioned before, the present study is mainly divided into three major phases. The first grounding pre-experimental phase that aimed to explore oral expression teachers' current teaching and assessment practices vis-à-vis the oral expression module. This phase was carried out using teachers' questionnaire that was designed and developed by the researcher.

The second phase was the experimental phase that encompassed the teaching intervention procedures. This phase started with pre-testing procedures of both control, and experimental groups. The two groups sat for the proficiency test. The end of this phase was devoted to a post proficiency test so as to examine the possible impact of the IPA implemented model.

For the third phase, the experimental group was subjected to an attitudes questionnaire in an attempt to understand the washback of the IPA model on their attitudes towards oral expression, in general, and the activities of the IPA, in particular.

Table 19 below, represents a summary of the phases, their objectives and the used methods of data collection in each phase.

Table 12.
Study Phases Description

Phase	Objective	Method	Tools
Grounding phase	The current assessment practices among the Algerian EFL teachers. Algerian EFL Teachers' perceptions and attitudes toward the current assessment practices and alternative assessment	Descriptive method	Teachers' questionnaire
Experimental phase	The use of the IPA model to enhance students' proficiency levels and rise their motivation in oral expression module.	Quasi-experimental "pre-post-test control group design	Pre-test Intervention (treatment) Post-test
The final phase	To examine the washback of IPA on students' attitudes	Descriptive method	Students' Questionnaire

3.5 Statistical Tools

The researcher made use of a various statistical measures and techniques. Both inferential and descriptive statistics were used. Descriptive statistics measures were used to describe participants' performance in the pre and post-tests, as well as, to describe their attitudes levels. While the inferential statistics measures were used to test this study's hypotheses, thus, they served to test the significance of the differences between experimental and control groups in the pre and post-tests. Inferential statistics were also used to test reliability and validity of the research tools.

Among the many measure of descriptive statistics, the mean (\bar{X}), standard deviation(S), and the skewness (Sk) were selected to describe the participants'

performance in the different tests conducted in this study. The following are the formulas of the three techniques:

The mean:

$$\bar{X} = \frac{\sum X}{N}$$

Where:

- $\sum X$: the sum of all the scores
- N: Sample size (number of participants)

The standard deviation:

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$$

Where:

- $\sum (X - \bar{X})^2$: the sum of squared deviations from the mean
- N: Sample size (number of participants)

The skewness:

$$Sk = \frac{\bar{X} - Mo}{S}$$

Where:

- \bar{X} : the sample mean
- Mo : the sample mode
- S: the standard deviation

The researcher made use of two types of T-test, namely; the paired sample T-test, to compare the experimental group performance in the pre and post testing, and the two independent equal groups to compare the experimental and control group performance in the pre and post testing. The formulas for the two tests are the following:

Independent equal samples T-test:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{S_1^2}{N_1} + \frac{S_2^2}{N_2}}}$$

Where:

- \bar{X}_1 : the mean of sample 1
- S_1 : the standard deviation of sample 1
- N_1 : the sample size of sample 1
- \bar{X}_2 : the mean of sample 2
- S_2 : the standard deviation of sample 2
- N_2 : the sample size in sample 2

Paired sample T-test:

$$t = \frac{\Sigma D}{\sqrt{\frac{n\Sigma D^2 - (\Sigma D)^2}{(n-1)}}}$$

Where:

- ΣD : the sum of the differences between scores in the pre-and post-tests
- n : sample size

Pearson correlation coefficient is also used to run different reliability and validity of the used tools, including, inter and intra rater reliability and concurrent validity. The formula of Pearson correlation coefficient (r) used in this study is the following:

$$r = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{\sqrt{[n(\Sigma x^2) - (\Sigma x)^2][n(\Sigma y^2) - (\Sigma y)^2]}}$$

Conclusion

Ultimately, this chapter presented all the methodological procedures and their underpinning rationales. It explained the inductive nature of this study that involved a Mixed-methods triangulation design, including both quantitative and qualitative methods of data collection and analysis.

Given the complexity of this study's variables, several methods and data collection tools were used. Two questionnaires were designed and administered in addition to an experiment with "pre-post-test control group design. The first questionnaire was a grounding pre-experimental phase that aimed to explore oral expression teachers' current teaching and assessment practices vis-à-vis the Oral Expression module. This phase was carried out using a teachers' questionnaire that was designed and developed by the researcher. The second questionnaire was administered to the students of the experimental group, and was subjected to explore their attitudes in an attempt to understand the washback of the IPA model on their attitudes towards oral expression in general and the activities of the IPA.

Reliability and validity of the data collection tools were examined using different methods in relation with the nature of the tool. Face validity was examined for the qualitative data collection tools. As to the quantitative tools, face and content validity were also examined. The Reliability, then, of the IPA assessment model was examined using two techniques including Cronbach's alpha (α) and the Guttman Split-half reliability analysis. The collected data from the tools were processed using computer software SPSS version 24 for Microsoft. Then, both qualitative and quantitative methods of data analysis were used. At the end of this chapter, a

description of the various statistical measures and techniques implemented in this study was provided.

To test the effectiveness of using the IPA in enhancing students' language proficiency, the researcher of this study opted for an intervention. Therefore, this chapter endeavours to provide a comprehensive explanation of the experimental phase that encompasses the intervention procedures. This phase started with pre-testing procedures of both control and experimental groups. The two groups sat for the proficiency test. The end of this phase was devoted to a post proficiency test so as to examine the potential impact of the implemented "IPA" model.

As it was highly emphasised in this chapter, the treatment, that is the IPA model, involved three phases constituted three separately established modes of communication "Interpretive, interpersonal and presentational". This stage constituted the body of this study's findings as it provided a rich data about the development of the IPA.

Chapter Four:
Data Analysis and Interpretation

Introduction

Before proceeding with the analysis of the data, obtained from the different data collection methods and testing this study's guiding hypothesis, a number of tests were conducted to ensure that the collection of the data is valid and reliable. This objective, has been achieved by two main levels of the study: The level of data collection methods, and the experimental procedures and measures.

Thus, the first steps in the data analysis chapter are devoted to the analysis of the used IPA model reliability and validity and the analysis of experiment groups homogeneity and normality of distribution of the targeted variables.

4.1 Normality distribution and homogeneity of the research sample

To ensure one of the conditions of the experimental design, which is normality distribution of language proficiency among the experimental and control groups, the Kolmogorov-Smirnov and Shapiro-Wilk normality tests were used. The data obtained from the analysis of normality among the experimental group levels of proficiency is represented on Table 20 below.

Table 13

Kolmogorov-Smirnov and Shapiro-Wilk normality tests (experimental group)

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	Df	Sig.	Statistic	df	Sig.
Experimental Group	,161	26	,082	,902	26	,079

a. Lilliefors Significance Correction

From table 20 above, examination of the statistical significance of both tests $p=0.82$ and $p=0.075$ that are more than $p > 0.05$ confirms that the data of these groups are normally distributed. The same test was conducted to the control group. Table 21 represents results of the same test.

Table14.

Kolmogorov-Smirnov and Shapiro-Wilk normality tests (experimental group)

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	Df	Sig.
Control group	,243	26	,061	,886	24	,081

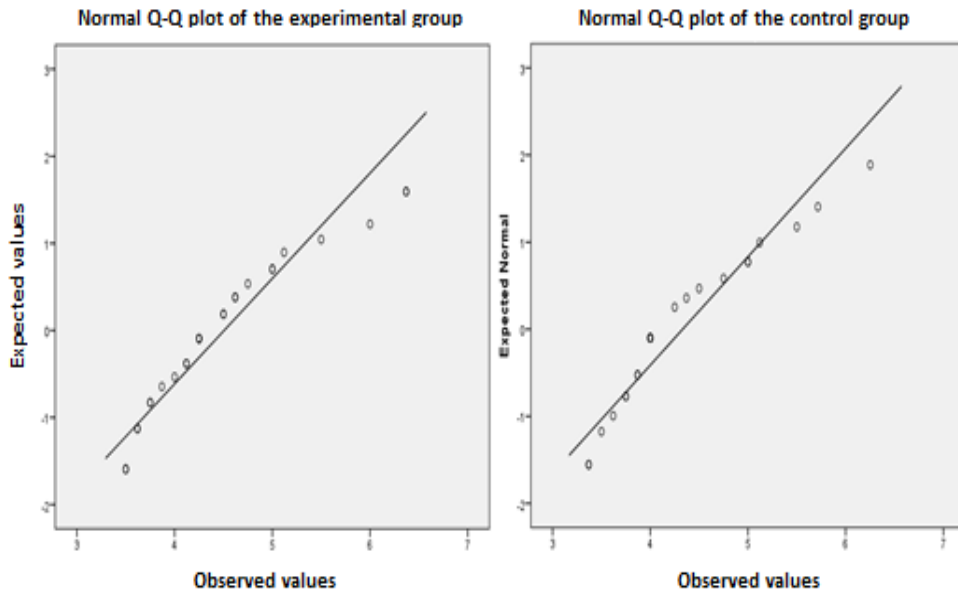
a. Lilliefors Significance Correction

Similarly, the detected values of the significance of the Kolmogorov-Smirnov and Shapiro-Wilk normality tests are more than 0.05 ($p > 0.05$), revealing the normality of distribution of proficiency scores among control group sample.

Moreover, the Q-Q scatterplots for normally distributed data generated by SPSS for both groups, the experimental group on the left and the control group on the right, confirmed the normality of the distribution due to the observed strong correlation between the expected and observed values of the quartiles.

Graph1

Graph of Normality Distribution Q–Q Plot



Another condition to ensure the use of the T-test is the homogeneity of the experimental and control group. Accordingly, both groups must have the same levels distribution. To test the homogeneity of distributions among both groups, the Levene's test of homogeneity of variance that tests the equality of variances. Table 22 below displays the results of this test.

Table 22.

Leven's Test of Homogeneity of Variances

	Levene's Test for Equality of Variances			
	Statistic	df1	df2	Sig.
Experimental –control	,000	1	48	,994

According to Table 25, the value of significance level that is $p= 0.994$, which is more than 0.05 ($p>0.05$), lead to the acceptance of the null hypothesis that states that there is no difference between the groups variances and then the conclusion that the groups have equal variances. Since the groups have equal variances, it can be said that they are homogeneous.

4.2 IPA proficiency test reliability and validity

Validity and reliability are two unavoidable qualities of measurement. To this end, the IPA test used in this study was subjected to validity and reliability tests during the piloting phase. The test was piloted with a group of 25 students in the piloting phase. The piloting sample were subjected to four IPA sessions. Two raters (teachers) were asked to rate the students' performance. The resulting data was processed to conduct a number of reliability and validity tests.

According to Hyland (2003), an assessment tool is reliable "If it measures consistently, both in terms of the same student on different occasions and the same task across different raters" (p.215). In regard to the IPA used in this study, the type of reliability that was examined was raters' reliability.

The piloting sample performances were scores by two different raters separately. The obtained scoring lists were used to run the inter-rater reliability statistically using correlation analysis. The same raters were asked to rate the same performance after a week of their first rating. This procedure was set to examine the intra-rater reliability.

The inter-rater and intra-rater reliability were run using intra-class correlation coefficient (ICC) of both single measures, which is an index for the

reliability of the ratings for one, typical, single rater, and average measures that is an index for the reliability of different raters averaged together. Since, from table 23, the values of ICC (0.938) for single measures and (0.899) for average measures are more than 0.7 ($ICC > 0.7$), we can say that there exists an excellent agreement between the two raters rating and one rater different ratings.

Table 15.

Intra-class Correlation Coefficient of Test Raters

	Intra-class Correlation ^b	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	.938	.993	.998	562.208	24	24	.000
Average Measures	.889	.996	.999	562.208	24	24	.000

Further, to examine concordances between the two rates ratings in the different rubrics of the analytical scoring grid of the IPA, Pearson Correlation coefficient (r) as used, thus correlation analysis was carried between the different scorings the raters on different components of the IPA scoring grid. The obtained values of rare displayed in table 24.

Table 16.*Reliability Analysis of the Raters' Scoring for the Scoring Scale*

		The interpretive (rater1)	The interpersonal (rater1)	The presentationa l (rater1)
The interpretive (rater2)	Pearson Correlation	.981**		
	Sig. (2-tailed)	.000		
	N	25		
The interpersonal (rater2)	Pearson Correlation		.988**	
	Sig. (2-tailed)		.000	
	N		25	
The presentational (rater2)	Pearson Correlation			.986**
	Sig. (2-tailed)			.000
	N			25

** . Correlation is significant at the 0.01 level (2-tailed).

From table 24 above, the values of r of the different components (the interpretive $r=0.981$, the interpersonal $r= 0.988$ and the presentational $r= .986$) are statistically significant in the level p , $p \leq 0.01$. The values of r also point to a strong correlation between the raters scorings assigned to each of the components of the IPA scoring scale.

The validity of the IPA was tested through the criterion validity method, exactly, through the use of concurrent validity. The latter is defined as “the degree to which scores on one test are related to scores on a similar, pre-existing test administered in the same time frame or to some other valid measure available at the same time” (Gay, 2012). Accordingly, this validity test was conducted

through comparing the students' performance in the IPA activities with their performance in the proficiency standardised test. Table 25 represents the results of the correlation analysis between the piloting sample levels in the IPA and the proficiency test.

Table 17.

Concurrent validity of the IPA test.

		Reading (PT)	Writing (PT)	Listening (PT)	Speaking (PT)
Reading (IPA)	Pearson Correlation	,451**			
	Sig. (2-tailed)	,001			
	N	25			
Writing (IPA)	Pearson Correlation		,310**		
	Sig. (2-tailed)		,000		
	N		25		
Listening (IPA)	Pearson Correlation			,480**	
	Sig. (2-tailed)			,000	
	N			25	
Speaking (IPA)	Pearson Correlation				,324**
	Sig. (2-tailed)				,018
	N				25

** . Correlation is significant at the 0.01 level (2-tailed).

From the correlation analysis table 25 above, the values of the validity coefficient are statistically significant at the level $p=0.01$. It is also noticed that the values range from 0,310, as a lowest value, to 0.451 as a highest value. These values represent acceptable to good values of the validity coefficient. Accordingly, we can say that the IPA used in this study is valid.

4.3 Analysis of Teachers' Questionnaire

Analysis of the questionnaire data is divided into two sections. The first section tackles the teachers' perceptions of and practices about teaching oral

expression and the second section probes for their perceptions and assessment practices vis-à-vis the oral expression module.

4.3.1 Teachers Perceptions and Practices of Oral Expression Teaching

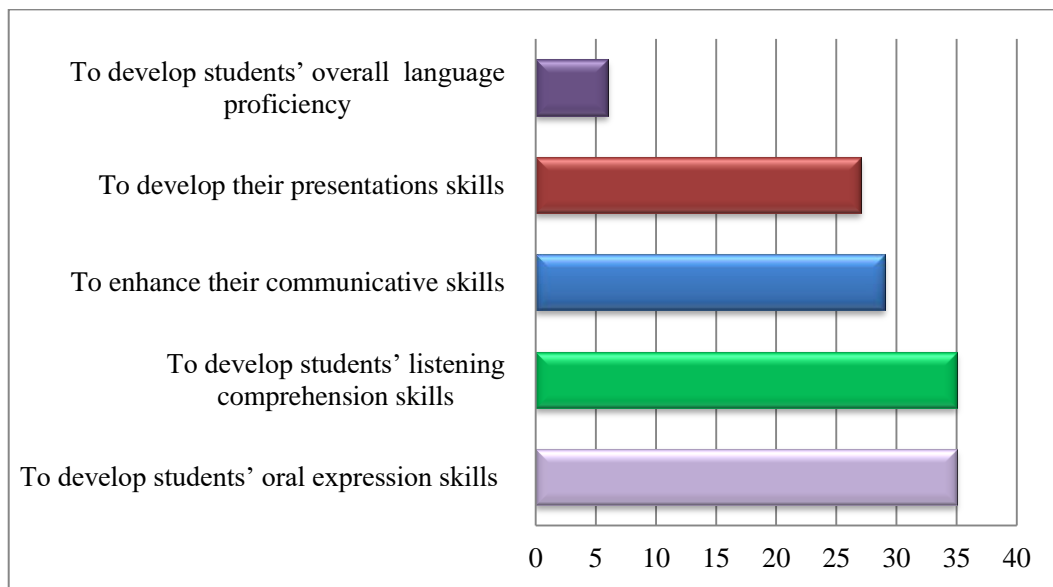
The first item in the second section of the questionnaire sought to detect the teachers' perceptions of about the objectives of teaching oral expression. The responses to item four are summarised in table 26 and graph 2.

Table 18.

Objectives of Teaching Oral Expression Module

Objectives	Freq	%
To develop learners' oral expression skills	35	100,0
To develop learners' listening comprehension skills	35	100,0
To enhance learners' communicative skills	27	77,14
To develop learners' presentations skills	21	60,00
To develop learners' overall language proficiency	12	34,28

From table 26, all teachers agreed on the first two objectives of teaching oral expression modules which are: developing the students' oral expression skills (100%) and developing the listening comprehension skills (100%). Among them, 77,14 % think that the written expression module's objective is to enhance learners' communication skills. As for developing communication skills, 27 teachers out of the total number agreed on it. Interestingly, only 12 teachers think that it serves to develop learners' overall language proficiency. This finding is clearly demonstrated in graph 2.

Graph 2*Objectives of teaching oral expression module graph*

The second rank item in the second section attempts to explore teachers' perceptions of the targeted skills in the oral expression module. Responses to item 5 are summarised in table 27 below.

Table 19.*Four Skills Importance Rank in Oral Expression Session*

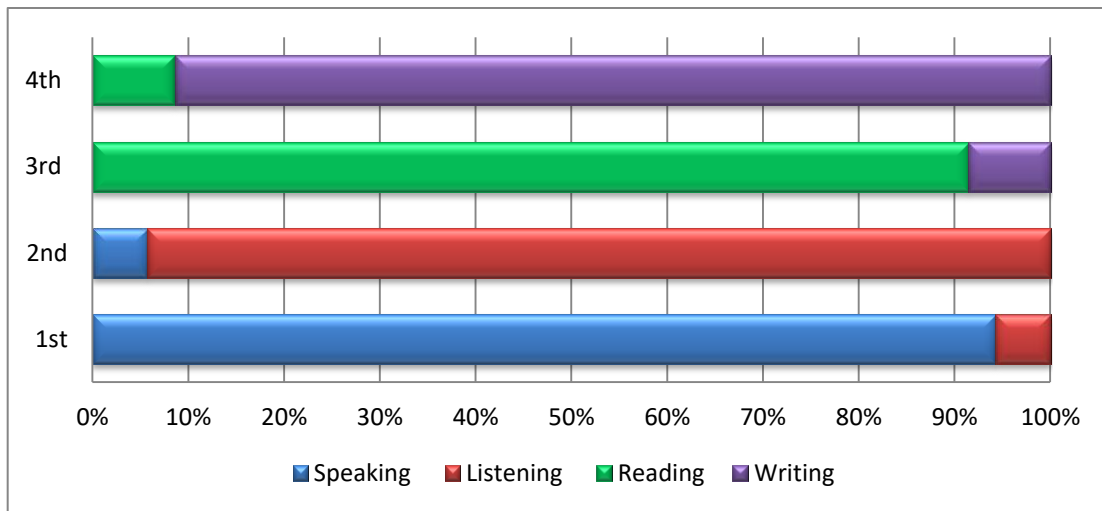
The skill	Rank	Freq	%
Speaking	1	33	94,28
Listening	2	33	94,28
Reading	3	32	91,42
Writing	4	32	91,42

From table 27, it is clearly illustrated that the speaking and the listening skills are ranked as the first two skills to gain higher focus among oral expression teachers. The reading skill was ranked as the third skill to be considered in the oral expression session while the writing skill received the last position with

91,42% of agreement that it is the least important skill in this session. Graph 3 clearly displays the different ranks.

Graph 3.

Four skills importance rank in oral expression session Graph



Item 6 data, asking the teachers about the most used teaching activities in the oral expression session, are summarized in table 28.

Table 20.

Oral Expression teaching activities

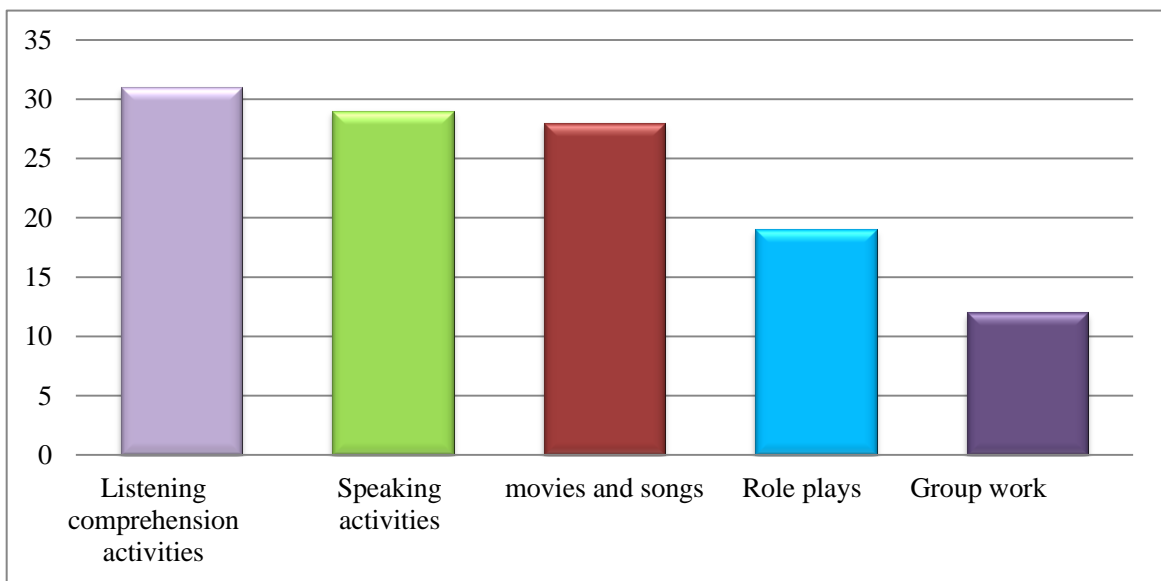
	Freq	%
Listening comprehension activities	31	88,57
Speaking activities	29	82,86
Movies and songs	28	80,00
Role plays	19	54,29
Group work	12	34,29

From table 28, it seems that the majority of teachers use the listening comprehension activities (88,57%). The Speaking activities were also widely used by the respondents as it received 82,86% of the responses. Equally, movies and songs were used by 80% of the teachers. The role plays were moderately used

by 54,29%. The group work was not widely used as only 34,29% of use. The teachers added other activities including, debates (5 teachers), which are part of the speaking activities, oral presentations (8 teachers) and reading circles (1 teacher). This finding is clearly illustrated through graph 4.

Graph 4

Oral Expression teaching activities graph



Responses to item 7, asking teachers about which activity is more enjoyed by students, are summarised in table 31 below. From the teachers' responses, movies and songs are the most enjoyed activities (94,29%).

Table 21.

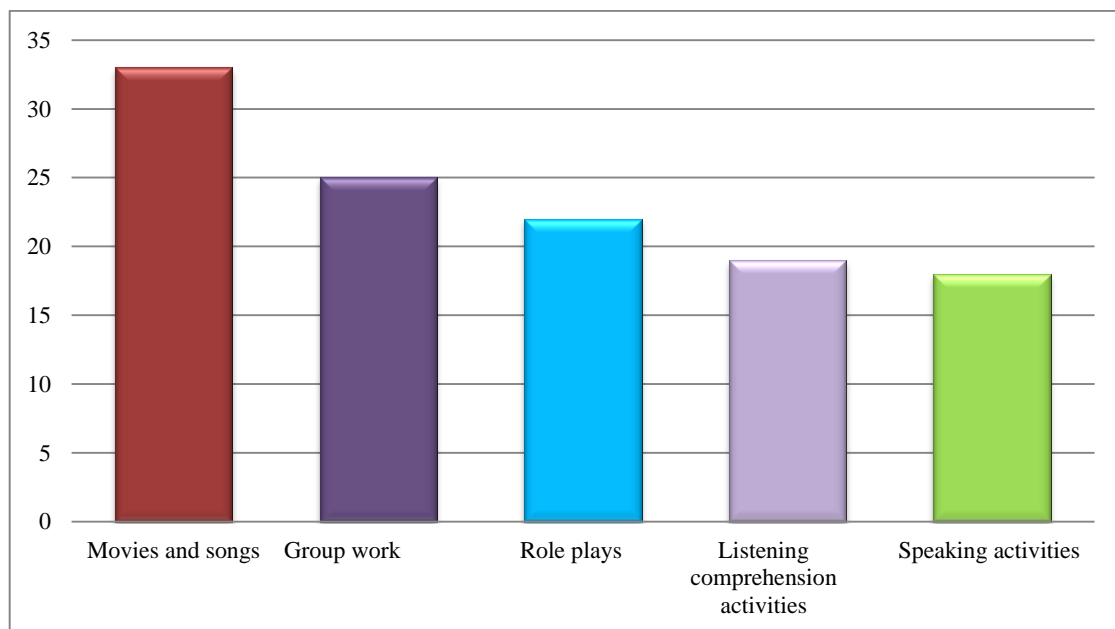
Students preferred oral expression teaching activities

	Freq	%
Movies and songs	33	94,29
Group work	25	71,43
Role plays	22	62,86
Listening comprehension activities	19	54,29
Speaking activities	18	51,43

Moreover, group work (71,43%) is viewed by teachers as another enjoyable activity in addition to role play (62,86%). The least enjoyed activities, according to teachers, were the listening comprehension activities (54,29%), and the speaking activities (51,43%). Graph 5 summarises the finding from item three.

Graph 5.

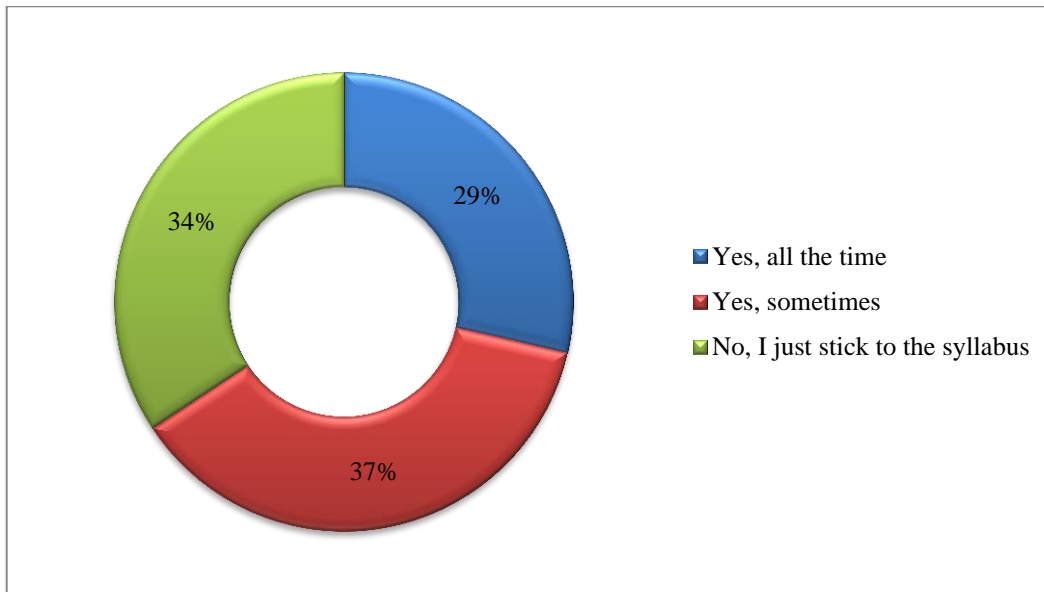
Most enjoyed oral expression teaching activities graph



Data from item eight, summarising teachers' responses about whether they allow their students pick up the topics and activities of their interest, are represented in graph 6.

Graph 6.

Students' participation in the selection of teaching content graph



From the pie chart display, it is confirmed that the majority of respondents (37%) do allow their students to participate in the section of the teaching activities and the contents, yet, not all the time. Only 29% do it all the time. A good number (34%) stick to the syllabus.

As for hindrances that they encounter when teaching the oral expression, explored by item 9, teachers' responses are summarised in table 30.

Table 22.

Hindrances faced by teachers when teaching oral expression

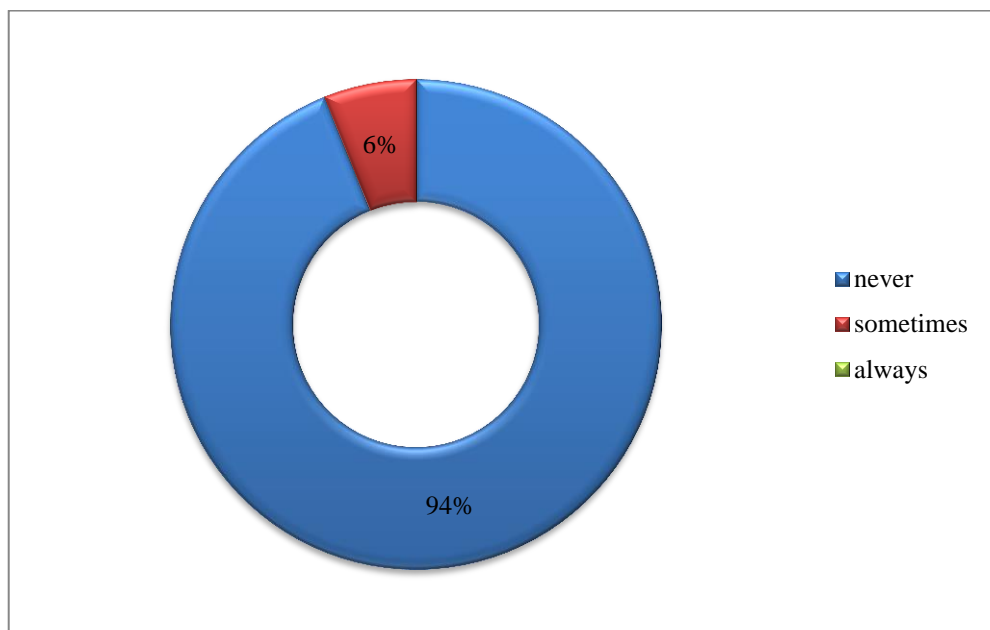
Hindrances	Freq	%
Weakness in language proficiency	33	94,29
Lack of motivation	28	80,00
Overcrowded classes	19	54,29
The type of topics and teaching activities	15	42,86
Logistics	10	28,57

According to table 30 data, it seems that among the encountered hindrances, weakness in language proficiency among the students constitutes the most problematic aspect (94,29%). Other factors as stated by the teachers include the lack of motivation among students (80%), overcrowded classes (54,29%). According to 42, 86% of the teachers the types of teaching activities do not help in motivating students. Logistics, including labs and ICTs tools were among the hindrances highlighted by teachers.

The last item (item 10), in the first section focus, was the exploration of teachers' perceptions about the use of the four skills in the oral expression session. Responses to this item revealed that the majority (94%), according to graph 7, do not integrate the four skills in the oral expression session. Only 6% use the four skills in the oral expression session.

Graph 7

The use of the four skills in the oral expression session graph



When asked through the sub-item, if item 10 about how teachers use the four skills in the oral expression session, the 6% of teachers explained that they use them as helping activities to prepare students for the speaking activities, or as homework to prepare for the coming sessions.

4.3.2 Teachers Perceptions of and Practices in the Oral Expression Assessment

The third section of the questionnaires is devoted to explore the teachers' perceptions of and practice about assessment in the oral expression classroom. The first item in this section probes for the assessment methods used by the participants. Data of item 11 are summarised in table 31.

Table 23.

Assessment Methods Used by Oral Expression Teachers

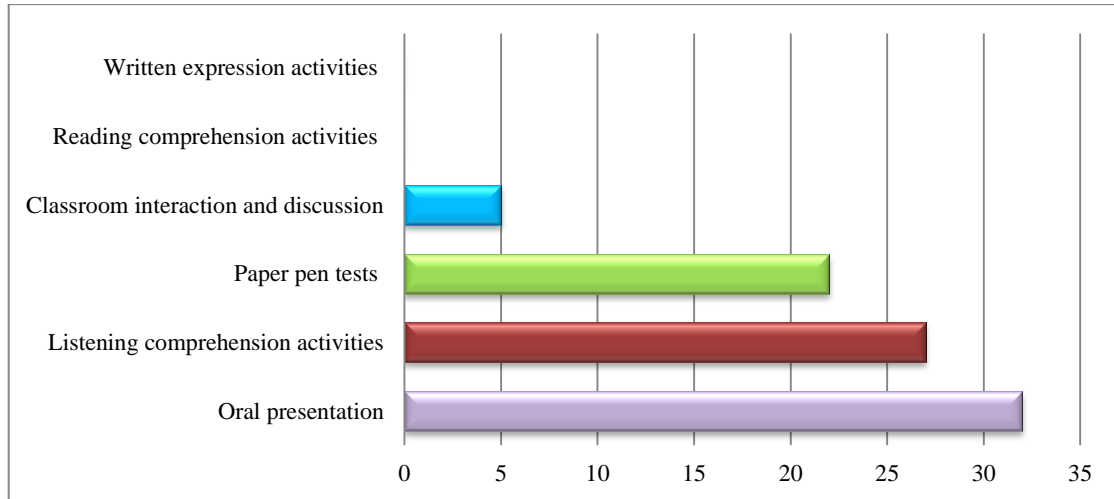
Assessment method	Freq	%
Oral presentation	32	91,43
Listening comprehension activities	27	77,14
Paper pen tests	22	62,86
Classroom interaction and discussion	5	14,29
Reading comprehension activities	0	0,00
Written expression activities	0	0,00

Apparently, in both Table 31 and Graph 8 below, the oral presentations constitute the top assessment methods (91,43%) used to assess students in the oral expression session, followed by the listening comprehension activities (77,14%), and paper pen tests (62,86%). Classroom interactions and discussion received only 14,29% of agreement, while both reading comprehension and written expression received no agreement on the possibility of

using them to assess oral expression skills. This finding is clearly displayed in graph 8 below.

Graph 8

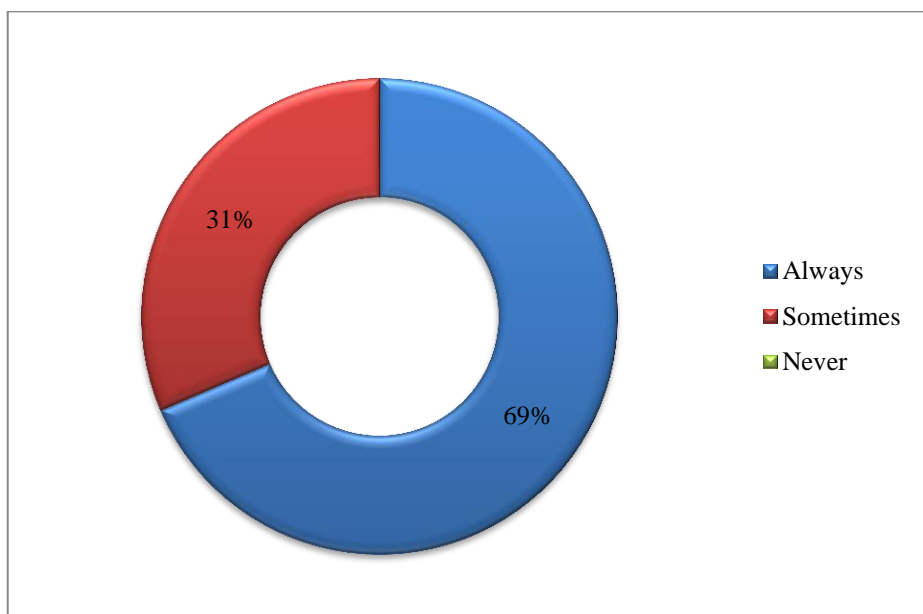
Assessment Methods Used by Oral Expression Teachers graph



Specifically, item 12 targets the frequency of teachers' feedback provision during the oral expression sessions. The responses lead to the data summarised in graph 9.

Graph 9.

Teachers' Feedback Provision Frequency graph



According to the graph's data, the teachers do always provide feedback to their students in all the sessions of oral expression. Only 31% of them do occasionally provide feedback to their students' performances.

Item 13 data, targeting the teachers' feedback type during oral expression sessions, are represented through table 34. The data reveal that the most used type of feedback among teachers is the oral feedback (97,14%). Only 17,14% of the respondents indicated that they use peers' feedback. Written feedback is rarely used with only two teachers confirming their use.

Table 24.

Teachers' Feedback Type

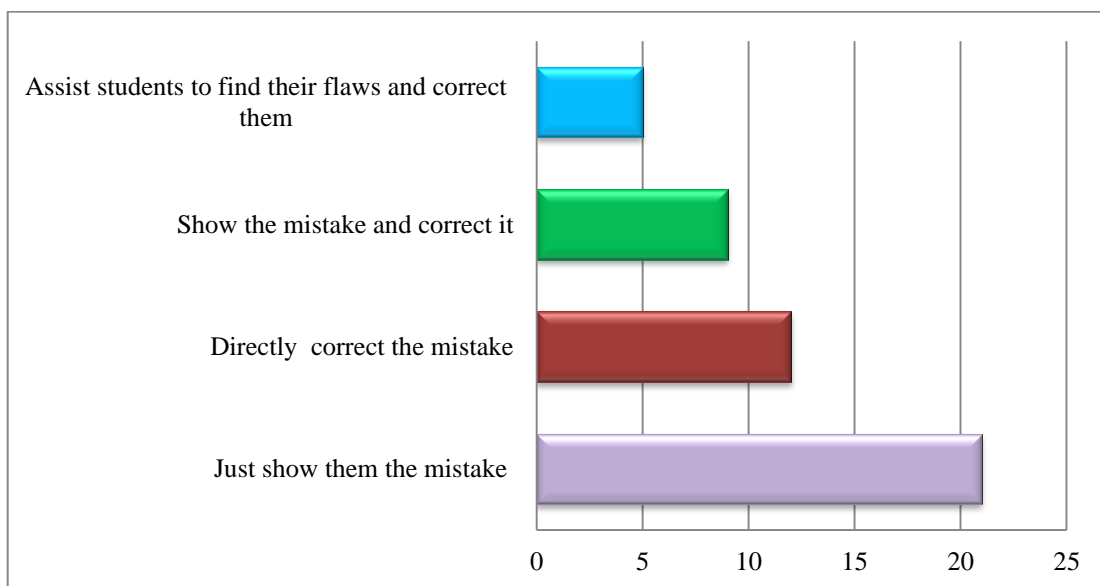
Feedback type	Freq	%
Oral feedback	34	97,14
Peers feedback	06	17,14
Written feedback	02	05,71

Keeping with feedback provision, further information was sought through item 14, which is set to ask the teachers about the method of feedback provision they usually use in their oral expression sessions. The responses are represented in table 33.

Table 25.*Methods of Feedback Provided by Oral Expression Teachers*

Methods	Freq	%
Directly correct the mistake	21	71,43
Show the mistake and correct it	12	51,43
Just show them the mistake	9	34,29
Assist students to find their flaws and correct them	05	14,29

From table 33, the most used method of feedback is through direct correction of the mistake (71,43%). 51,43% of the respondents indicated that they do not only show the mistakes, but they also correct them, while 34,29% of them revealed that they just show the mistakes. Graph 10 displays findings from this item.

Graph 10.*Methods of Feedback Provided by Oral Expression Teachers graph*

Further, item 15 focuses on the types of mistakes teachers target when providing feedback. Data from item 15 are demonstrated in table 34.

Table 26.

Types of mistakes targeted by teachers' feedback

	FREQ	%
Pronunciation mistake	33	94,29
Grammatical mistakes	32	91,43
L1 transfer	15	42,85

The data displayed in table 36, show that most teachers correct the pronunciation (94,29%) and grammar (91,43%) mistakes. They also show that only 42,85% target the L1 transfer related mistakes.

As for the difficulties teachers face when assessing students' performance in the oral expression session, item 16 is devoted to detect these difficulties. Data of item 16 are thematically classified in table 35.

Table 27.

Difficulties Faced by Teachers when Assessing Oral Expression

	Freq	%
Overcrowded classes	33	94,29
The limited number of allotted sessions	34	97,14
The weak level of students	32	91,43

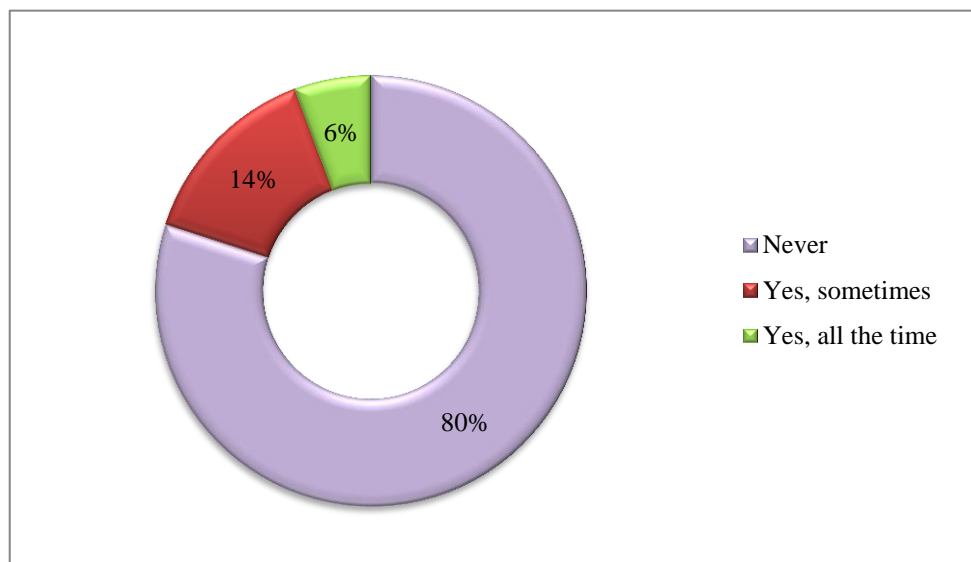
From table 35 data, all the teachers agreed on the most difficulties they encounter when assessing oral expression. The difficulties involve the number of

students per group, which received 94,29% of agreement; the limited number of allotted sessions, which received 97,14% of agreement and the weak levels of proficiency among the students which received 91,43% of agreement.

Initially, item 17 targets an important aspect in this study, which is the integration of the four skills in the process of evaluation of students' overall performance in oral expression sessions. The findings of this item are illustrated through graph 11.

Graph 11.

The Use of the Four Skills in oral expression assessment graph

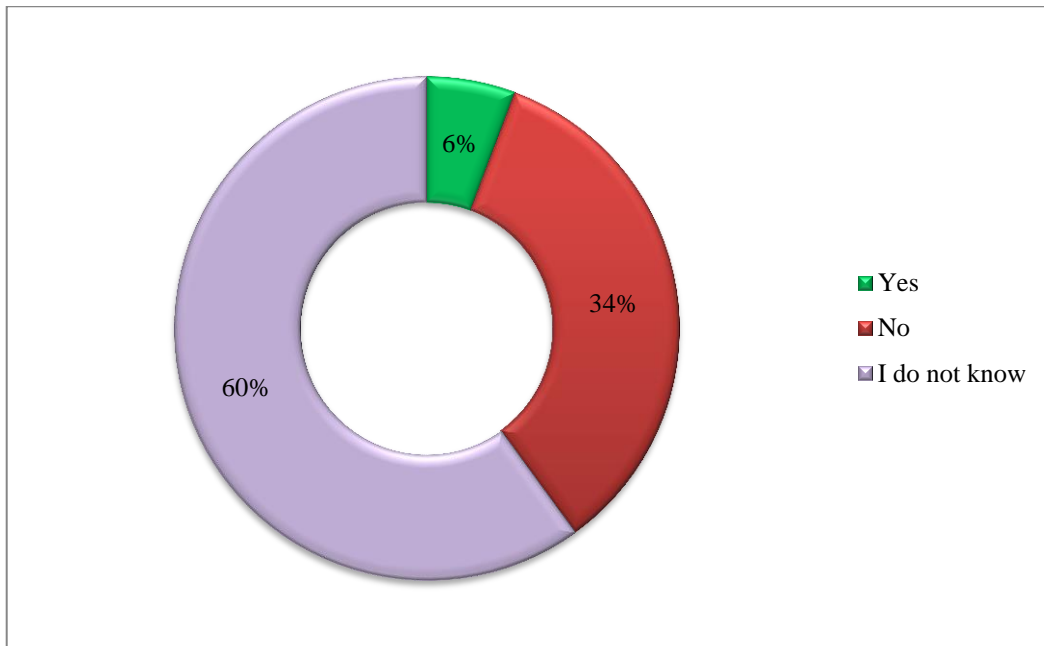


Clearly, 80% of the respondents confirmed that they never use the four skills in their assessment practices. Only 14% of the respondents confirmed that they occasionally use the four skills as part of assessment. Only 6% of the respondents indicated that they always use the four skills.

To probe teachers' perceptions about the efficiency of the integration of the four skills in the evaluation process in the oral expression session teachers were asked to respond to item 18. Graph 12 exhibitions Data from item 18.

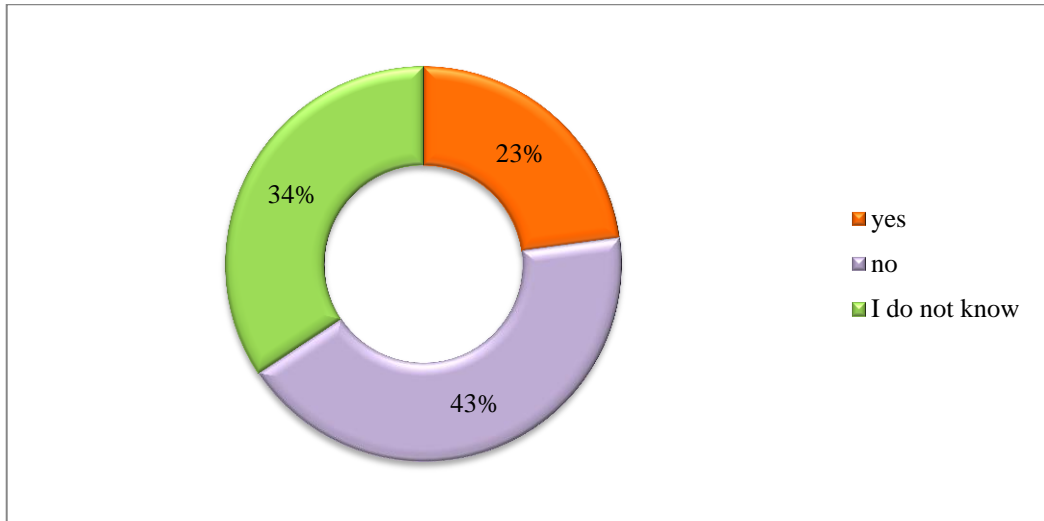
Graph 12.

Efficiency of the Four Skills Integration graph

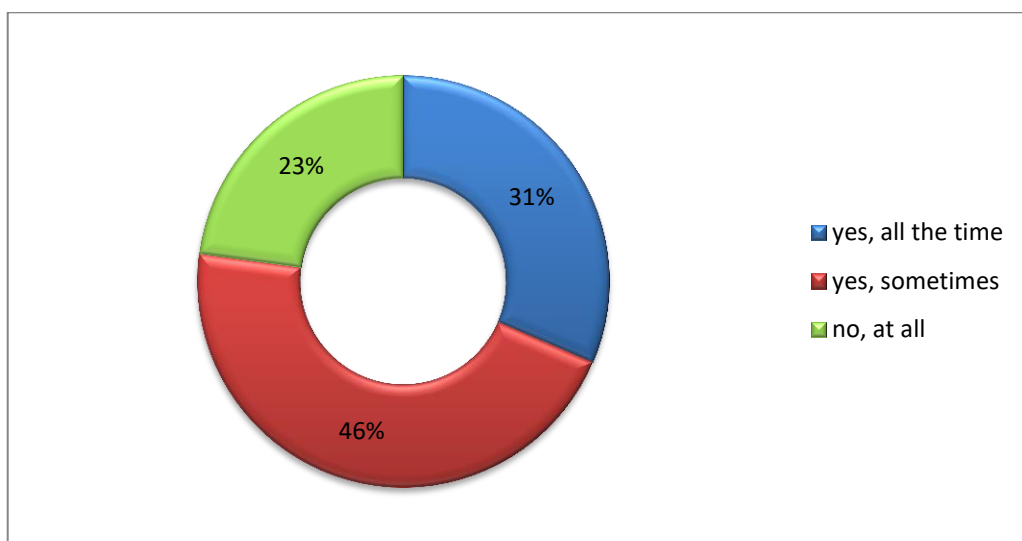


From the graph, not all teachers were sure about the efficiency of integrating the four skills in the evaluation process. Yet, 34% of the participants think that there is no need to integrate the four skills. Only 6% think that skills integration in the evaluation process.

Most teachers did not agree with item 19 that targets the use of traditional methods of assessment as main tools to assess students' performance. According to graph 13 below, most of the respondents think that the use of traditional methods of assessment is sufficient, while 23% believe in the opposite. A good number of respondents 34% have no idea about the efficiency of traditional assessment methods.

Graph 13.*The Use of Traditional Methods graph*

When asked, in item 20, about the use of alternative methods of assessment, 46% of teachers confirmed that they sometimes use alternative assessment methods in oral expression, while 31% confirmed that they always use them all the time. However, only 23% confirmed they never use these alternative methods in their assessment practices.

Graph 14.*The use of Alternative Assessment Methods graph*

The final data of the questionnaire were related to the last sub-item of item 20, which probes for the alternative methods used by oral expression teachers. The responses indicated the use of: group projects, role plays, ... and so on.

4.4 Analysis of the Experimentation Data

The data obtained from the experimental phase were collected from the pre- and post- testing stages. All the comparisons were run using the T-test for independent samples and T-test for paired samples. The difference analysis was carried to test this study's hypotheses. Thus, the order of proceeding sections is relative to the hypotheses set by the researcher.

4.4.1 Data analysis of the first hypothesis

The first Null hypothesis in this study belongs to the pre-testing phase stating that:

- There is no statistically significant difference between control and experimental groups' performance in the proficiency pre- test.

To test the null hypothesis both experimental and control group performances in the proficiency test before the treatment were subject to independent samples T-test. The data obtained from the first test are displayed in table 36.

Table 28.*Independent groups T-test between experimental and control groups pre-test performance*

		Levene's Test for Equality of Variances		t-test for Equality of Means		
		F	Sig.	t	df	Sig. (2-tailed)
Experimental -Control	Equal variances assumed	,021	,994	,763	48	,449
	Equal variances not assumed			,763	47,86	,449

Given the p value ($p=0.994$) of the Levene test of equality of variance, which is ($p > 0.05$), indicating the existence of equal variances of the groups, which in turn point to the use of the T-test in case of equal variance assumed. Accordingly, the t value ($t=0.763$) is not statistically significant due to the fact that the related p value ($p=0,449$) is more than 0.05 ($p > 0.05$). This implies the acceptance of the null hypothesis stating that there are no statistically significant differences between experimental and control groups' performances in the pre-test.

4.4.1.1 The description of experimental and control groups proficiency levels in the pre-test

Before conducting the second phase of difference analysis, it is important to understand the experimental and control groups' global performance in the proficiency test in order to gain a better understanding of the strengths and weaknesses in the study sample, as well as having a clear vision on their level before the integration of the IPA model.

The first set of descriptives are summarised on table 37 below. From this table 39, the mean value of the global scores obtained by the experimental sample ($\bar{X}=4.51$), which is a B1 level reveals that the students proficiency levels are intermediate. Examining the standard deviation, which is so small (Std.D =0.83), it is assumed that most of the scores are close to the mean confirming that all the participants hold the same level. This conclusion is confirmed by the small range value (R=2.87) indicating a slight dispersion between the sample scores.

Table 29.

Descriptives of the Experimental group performance in the pre-test

	N	Range	Min	Max	Mean	Std. Deviation	Skewness Statistic	Std. Error
Exp.Group	26	2,87	3,50	6,37	4,51	,83	,98	,46
Valid N	26							

Further analysis of the classifications of the sample assigned proficiency levels are provided in table 38 and graph 15.

Table 30.

Distribution of experimental groups pre-test proficiency levels

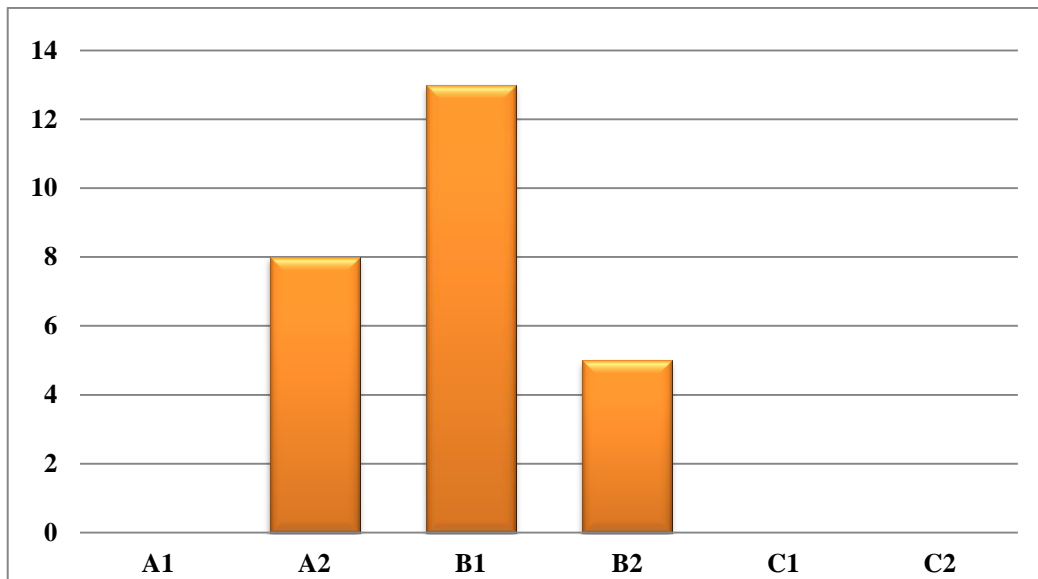
	Frequency	Percent	Cumulative %
A2	8	30,7	64,0
B1	13	50,0	90,0
B2	5	19,3	100,0
Total	26	100,0	

From table 38, it seems that half of the respondents (50%) of the sample hold the B1 level, which is an intermediate level. Among the participants 30.7%

(eight students) are basic users with the A2 level, while five students are B1, which are close to intermediate independent users of English. Graph 15 below shows clearly the distribution of levels among the experimental group members.

Graph 15.

The experimental group proficiency pre-test levels graph



Supplementary analysis of the four skills targeted by the Proficiency test was conducted to examine the participants' strengths and weaknesses. Table 39 below summarises the descriptives of the four skills performance.

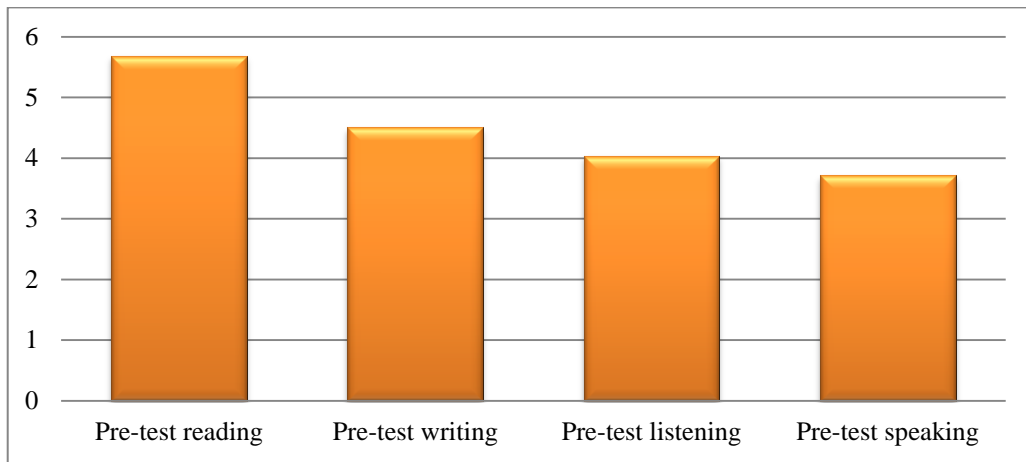
Table 31.*Experimental group pre-test performance in the four skills*

	N	Range	Min	Max	Mean	Std. Deviation	Skewness	
							Stat	Std. Error
Pre-test reading	26	4,00	4,00	8,00	5,76	,87	,61	,456
Pre-test writing	26	3,50	3,00	6,50	4,50	1,08	,30	,456
Pre-test listening	26	3,00	3,00	6,00	4,04	1,05	,59	,456
Pre-test speaking	26	3,00	2,50	5,50	3,71	,88	,79	,456
Valid N	26							

Apparently, from table 39, the highest mean is the reading skill mean ($\bar{X}=5.76$) with the Std.D = 0.87, followed by the writing skill mean ($\bar{X}=4.5$, Std.D=1.08), then the listening skill mean ($\bar{X}=4.04$, Std.D=1.05), and the speaking skill mean ($\bar{X}=3.71$, Std.D=0.88). These findings reveal that the experimental group's highest performance was in the reading comprehension activities and their lowest performance was in the speaking skill while both the writing and listening were average performance. These findings are clearly demonstrated on graph 16 below.

Graph 16.

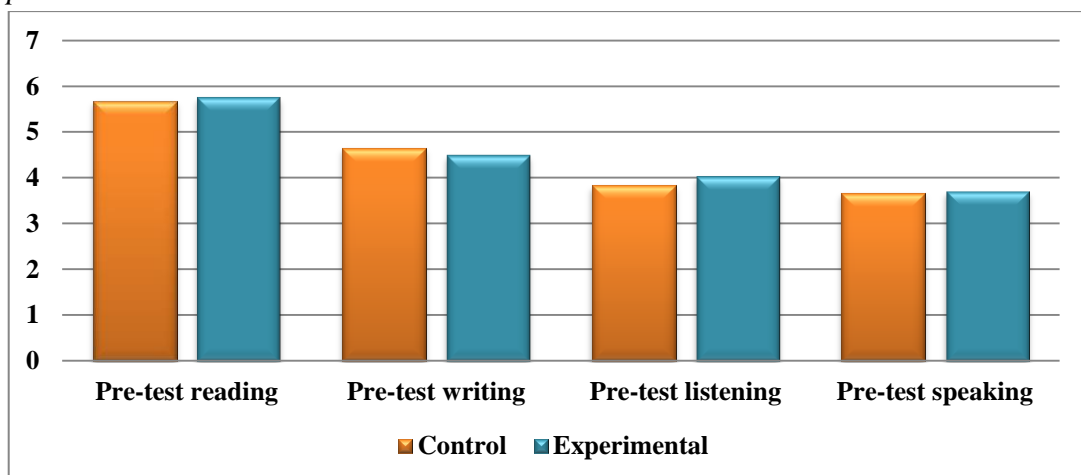
Experimental groups' four skills performance graph



As for the control group performance, and since it was confirmed in section 2, in this chapter, the experimental and control groups are homogeneous, including the same distribution of proficiency among the samples. There was no need to present the descriptives of the control. However, the following graphs serve to demonstrate the similarities of distributions of the proficiency among both samples. Graph 17, displays clearly the difference between the experimental and control groups' four skills performances.

Graph 17.

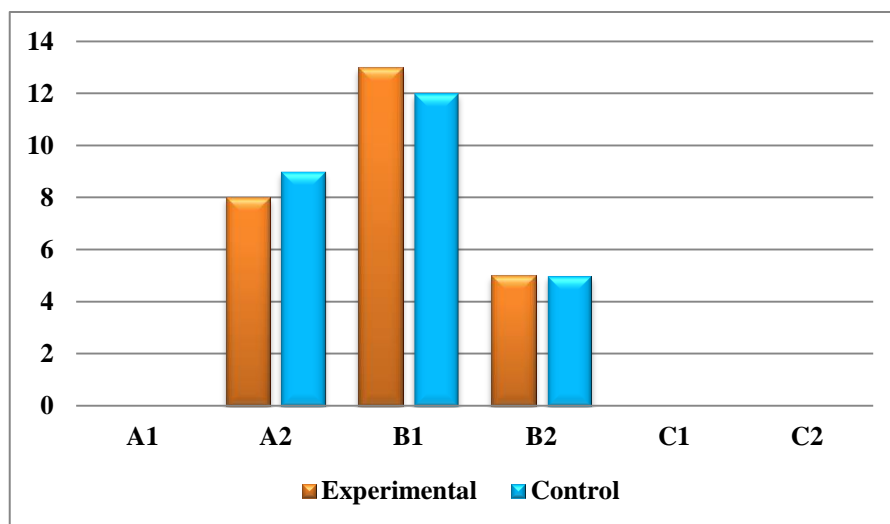
Differences between experimental and control groups four skill performance in the pre-test graph



It shows clearly on graph 17 between the two groups that the four skills performance are nearly identical with slight unmentioned differences in the reading, writing and listening skills. These findings confirm the difference analysis first conducted that accept the null hypothesis which, in turn, confirms the equality of performance among the two groups. Further evidence of equal performance among the two groups is validated by graph 18.

Graph 18.

Pre-test proficiency levels distribution among experimental and control groups graph



Clearly, the distribution of levels in the control group is similar to the control group, hence, confirming again the null hypothesis and rejecting the existence of differences in performance between the two groups.

4.4.2 Data analysis of the second hypothesis

The second hypothesis in this study is relevant to the experimental phase:

- There is no statistically significant difference between the experimental and control groups' performance in the proficiency post-test.

To test the second null hypothesis, the independent T-test was used. Table 40 represents the generated data from the difference analysis between control and experimental groups' performance in the post test.

Table 32.

Independent unequal groups T-test between the experimental and control groups pre-test performance

		Levene's Test for Equality of Variances		t-test for Equality of Means		
		F	Sig.	t	df	Sig. (2- tailed)
Experimental -Control	Equal variances assumed	1,88	,176	3,11	48	,003
	Equal variances not assumed			3,15	46,55	,003

The value of p ($p=0.176$, $p>0.05$) relevant to Levene's test of groups variances equality, which reveals that both groups have equal variances leads the analysis to read the values of T-test in the case of assumed equal variances. Accordingly, the t value ($t=3.11$) is statistically significant since $p=0.003$ is less than 0.05 ($p<0.05$). Consequently, and since $p<0.05$, the null hypothesis is rejected and the alternative one is accepted concluding that there exists a significant difference between experimental, and control groups performance in the post-test.

4.4.2.1 The description of experimental and control groups proficiency levels in the pre-test

To closely examine the detected difference between the experimental and control groups in the post-test, comparisons were made between the two groups' performances. Descriptives of both groups are summarised in table 41.

Table 33.

Descriptives of experimental and control groups post-test performances

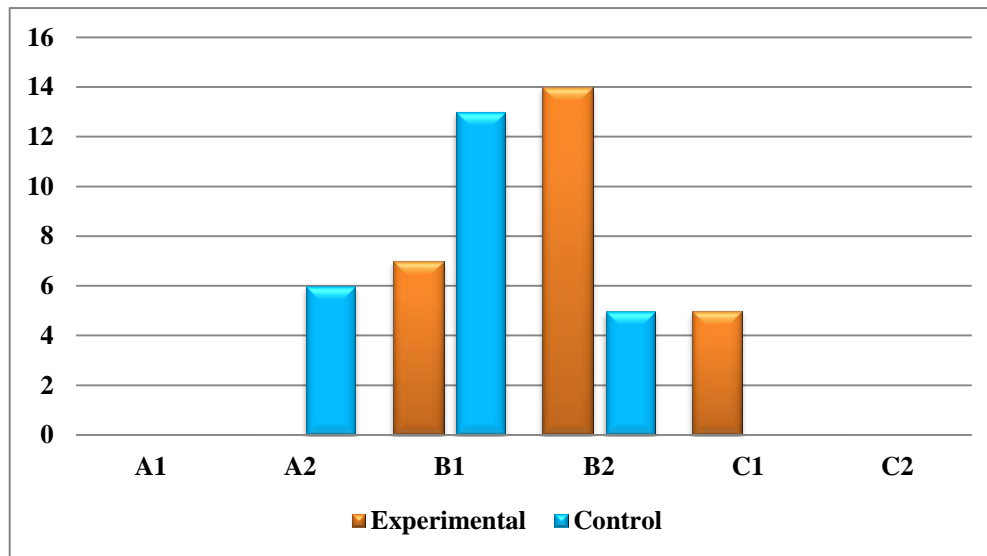
	group	N	Mean	Std. Deviation	Std. Error Mean
Post- testing	Experimental group	26	5,08	,67	,17
	Control group	26	4,07	,87	,13

The difference is the two groups means values is equal to 1,01. A closer analysis of the two groups' means values shows that the experimental (with $\bar{X}=5,08$, Std.D=0.67) group performed better than the control group (with $\bar{X}=4.07$ and Std.D=0.87). Thus we can say that the experimental group performed better than the control groups in the post-test.

To clearly demonstrate the differences, graphs 5 and 6 were generated. The first comparison was between the two groups distribution of the proficiency levels. Graph 19 represents the differences in the proficiency levels distributions.

Graph 19.

Post-test Proficiency Levels' Distribution among the Experimental and Control Group graph

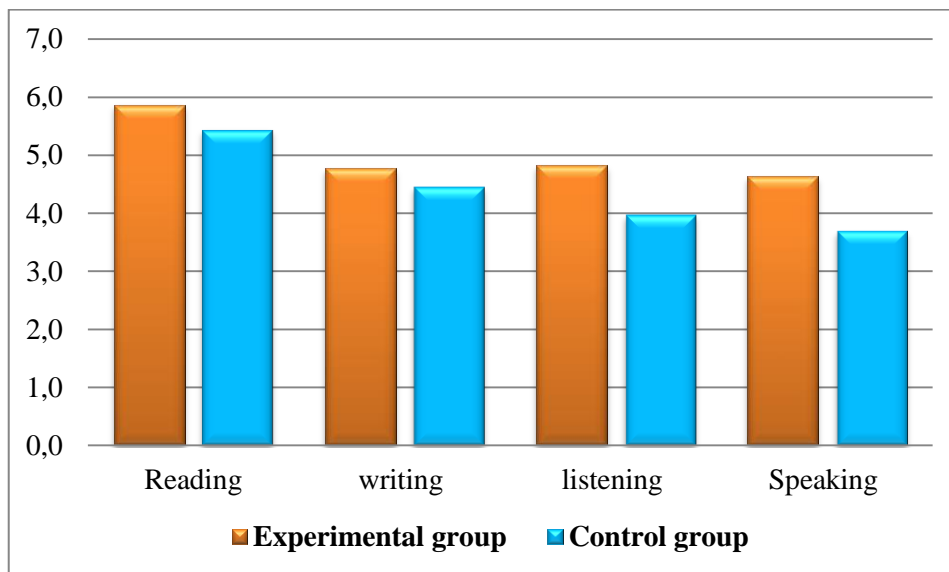


Undoubtedly, the bars of different levels, displayed in graph 05, show the unblemished difference between the proficiency levels distributions. The experimental group new distribution includes a C1 advanced level shift. Additionally, the absence of A2 beginner level, and an increase in B2 (close to the advanced level) appeared, in addition to a decrease in the B1 intermediate level. In contrast to the experimental group, the control group B1 is higher, and the B2 is lower and that the A2 is only found among the control group.

A comparison of the four skills was also consulted to examine the differences in performance among the two groups. Graph 20 exhibits the differences in the four skills performance in the post-test.

Graph 20.

The differences in the four skills performance in the post-test graph



Similarly, from graph 20, a notable difference is observed in the four skills. Remarkably, the difference is larger in the listening and speaking skills to the experimental group.

4.4.3 Data Analysis of the Third Hypothesis

The third hypothesis of this study states:

- There is no statistically significant difference between the experimental group performances in the pre and post-test.

The above hypothesis was tested through the use of Paired Samples T-test to compare between the experimental group pre and post-test performance. Table 42 represents the data of the Paired Samples T-test.

Table 34.

Paired sample T-test between Experimental group pre and post-test performance

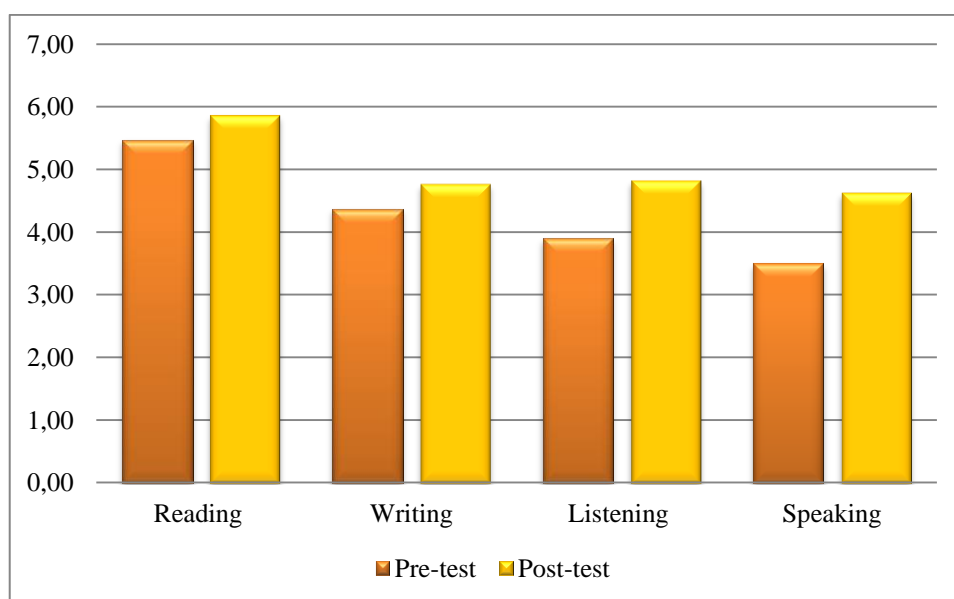
				Paired Differences			
		Mean	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
Pair 1	Pre-test	4,50	,83	,16	-7,80	25	,000
	Post-test	5,08	,87	,17			

Interestingly, a statistically significant difference between the experimental performances in the pre- and post-test is detected with a t value $t=7,80$ and p value $p=0.000$ which is less than 0.05 ($p<0.05$). Thus, the null hypothesis is rejected and the alternative one is accepted stating that there is a statistically significant difference between experimental group performances in the pre- and post-test.

To examine the direction of the difference and whether this difference implies an improvement in the experimental group performance or not, graph 21 was generated.

Graph 21.

Differences in the experimental group pre- and post-test performance graph

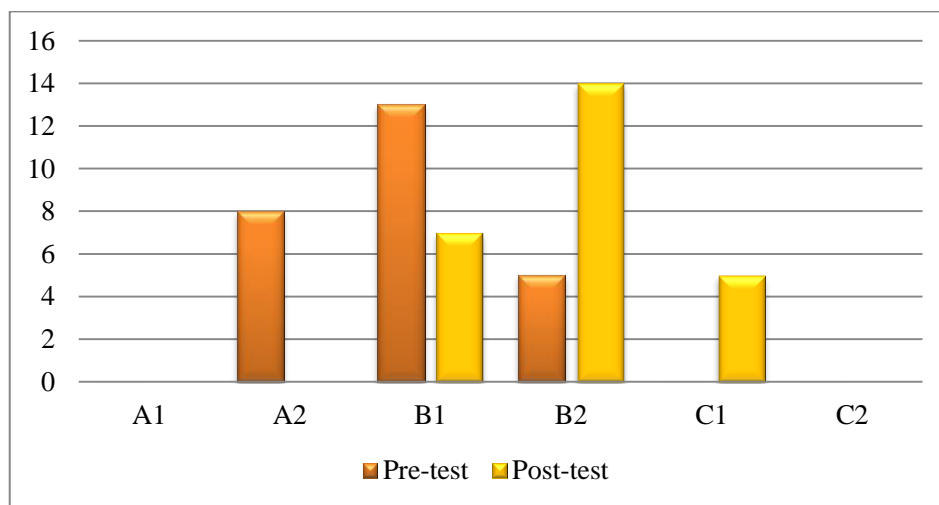


Evidently, as displayed in graph 21, the experimental group performed better in the post test. This finding is evidenced by the noticed increase in the post test bars compared to the pre-test ones.

The difference in proficiency levels distribution was also examined. Graph 22 shows the differences in the experimental group levels of proficiency in the pre and post-tests.

Graph 22.

Differences in the pre-and post-test proficiency levels distribution graph



The difference in performance is also confirmed by graph 22 bars. A new distribution in the experimental group is detected, wherein level C1 emerged as a new classification. A decrease in B1 level and an increase in B2 level are detected. These findings reveal a clear improvement in the experimental group's performance after the integration of the IPA.

4.4.4 Data Analysis of the Fourth Hypothesis

The last difference analysis was between the control group performance in the pre- and post-test to test the fourth null hypothesis stating:

- There is no statistically significant difference between control group performances in the pre- and post-test.

Analysis of the control group performance was necessary to ensure the internal validity of the experiment. Analysis of the difference in the control group performance in the pre-and post-test, without being subject to treatment, was conducted through the use of paired- sample t-test. The data obtained are summarised in table 43.

Table 35.

Paired sample T-test between Control group pre and post-test performance

			Paired Differences				
		Std.	Std. Error	Sig.			
		Mean	Deviation	Mean	t	df	(2-tailed)
Pair 2	Pre-test	4,33	,80	,16	-1,15	23	,263
	Post-test	4,39	,67	,13			

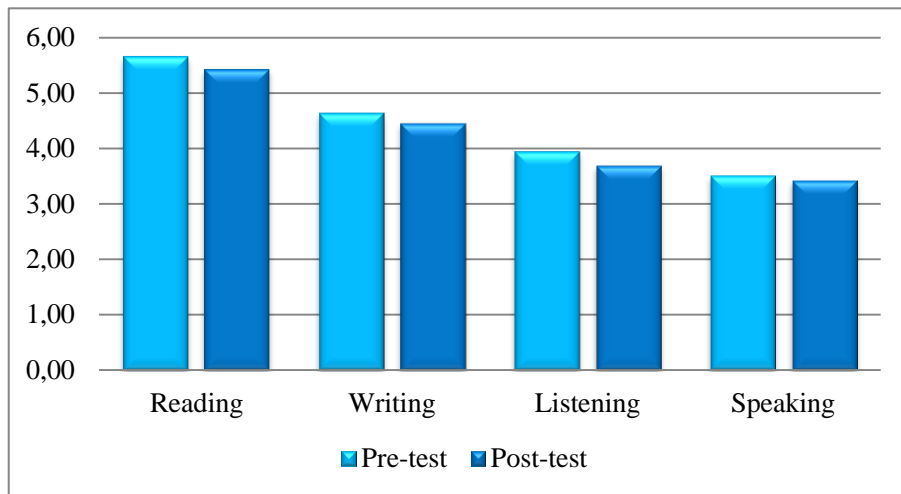
From table 43, the value of $p=0,263$ is more than $0,05$ ($p>0,05$), thus revealing the absence of statistical significance. Accordingly, the null hypothesis is accepted confirming the absence of statistically significant difference between the control group pre- and post-test performances in the proficiency test.

Graph 23 demonstrates the performances of the control group in the pre- and post-test. From the graph, it seems that there is no real change between the performances in the two phases. Yet, a slight decrease is detected in the four skills in the post-test. This decrease is clearly noticed in the reading, writing, and

listening bars. Thus, it is concluded here that there is no improvement in the control group performance in the proficiency test in the post-testing phase.

Graph 23.

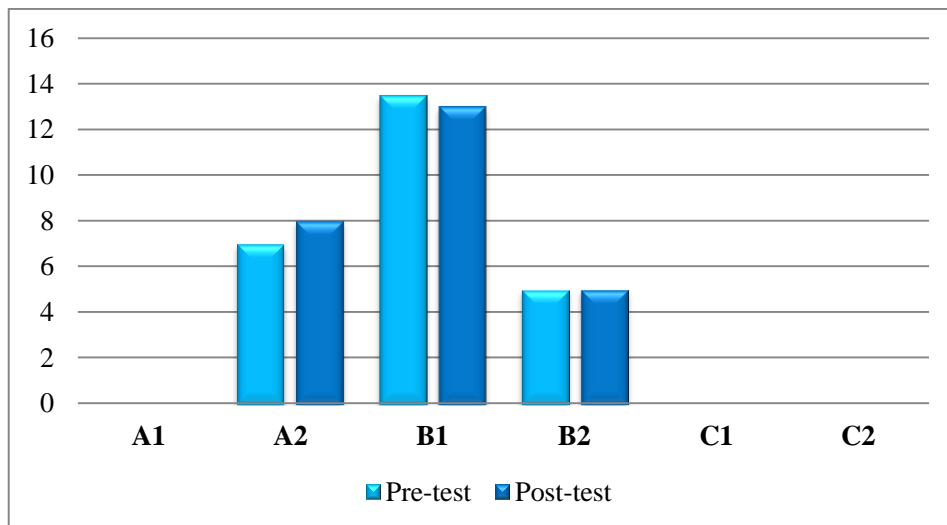
Differences in control group pre- and post-test performance graph



Similar to the experimental group, differences in proficiency levels distribution in the pre and post- tests. Graph 24 displays the differences if found.

Graph 24.

Differences in the pre-and post-test proficiency levels distribution among control group graph



In light of the displayed information on graph 10, it seems that there is no improvement in the proficiency levels among the control group. The levels are almost similar, except for an increase in the A2 level in the post-testing.

To conclude the experiment data analysis, the findings from the analysis of the four sub-hypothesis revealed the acceptance of the first null hypothesis, stating that there is no statistically significant difference between the experimental and control group in the pre-testing phase, and the acceptance of the second null hypothesis, stating that there exists a statistically significant difference between experimental and control groups performances in the post-testing. The third null hypothesis was rejected stating that there is no statistically significant difference between the experimental group performance in the pre- and post- testing phases. The fourth sub-hypothesis was confirmed stating that there is no statistically significant difference between the control group performance in the pre- and post-testing phases.

Thus from these findings, it is concluded that the main hypothesis is accepted confirming that the implementation of IPA assessment model has contributed to the promotion of the proficiency levels among the participants.

4.5 Students' Questionnaire Data Analysis

The students' questionnaire is designed to test the washback of the implemented IPA assessment model on students' attitudes and motivation towards teaching, learning, and assessment in oral expression. The analysis of students' questionnaire data is divided into two sections. The first section detects students' attitudes towards the oral expression session after the integration of the

IPA while the second section seeks to understand students' attitudes towards the IPA activities and outcomes

4.5.1 Students Attitudes towards Oral Expression New Sessions

The analysis of the Likert items of the first section was carried using the weighted average intervals method indicated in table 44 below.

Table 36.

Weighted mean intervals

Mean value	Level
[1- 1,8[Very low agreement
[1,8 – 2,6[Low agreement
[2,6 – 3.4 [Moderate agreement
[3,4 – 4,2[High agreement
[4,2 - 5]	Very high agreement

Thus, the comparison of the obtained means from students' responses to questionnaire to the weighted mean values served to reveal information about the levels of their attitudes towards oral expression module before and after the integration of IPA model of assessment. Data obtained from the first section of the questionnaire are summarised in table 45.

Table 37.*Students' attitudes towards the oral expression module*

Item	Statement	Mean	Std.D	Order
8	Oral expression sessions helped me develop my ideas	4,81	0,85	1
7	Oral expression session helped me improve my vocabulary	4,79	0,75	2
5	Oral expression second semester session is so motivating	4,77	0,82	3
9	Oral expression teacher cares about enhancing my performance for the better	4,50	0,52	4
1	Oral expression second semester session is more enjoyable	4,25	0,72	5
10	I do my best to show my teacher and classmates that I learnt from mistakes	4,25	0,75	6
6	Reading and writing helped me express my ideas orally	4,19	0,75	7
3	Oral expression second semester session is more beneficial and informative	4,16	0,80	8
11	Oral expression session helped me develop my communication skills	3,95	0,86	9
2	Assessment in the second semester is not stressful	3,88	0,75	10
4	Oral expression first semester session is more enjoyable	3,79	0,77	11

According to the mean values, the items are reordered downward from the highest mean to the lowest one. Examining the means of the items, it seems that items 8, 7, 5, 9, and 10 means ranging from $\bar{X}=4,81$, as the highest value, to $\bar{X}=3,79$, as the lowest value, confirms that they all belong to the higher agreement interval. In addition to this, the values of the standard deviation ranging from 0,85 to 0,52 points to the narrow spread of

students' responses from the mean, thus confirming that the majority of students hold the same levels of agreement.

Accordingly, the data summarised in table 45 above confirms the highly positive attitudes towards the oral expression session as all the students highly agree that the oral expression helps them improve their ideas (item 8), and their vocabulary (item 7). They also highly confirm that the oral expression session is more enjoyable (item 1), and motivating (item 5). The students hold highly positive attitudes towards the teacher' feedback (item 9), and peers' feedback (item 10).

Similarly, examining the mean values of items 6, 3, 11 and 2 means, ranging from $\bar{X}=4,19$ to $X=3,88$ that belong to the high agreement interval, demonstrates the students positive attitudes. The values of Std. D, ranging from 0,86 to 0,75, confirm that the majority of students' responses are close to the mean and not far from general level of agreement.

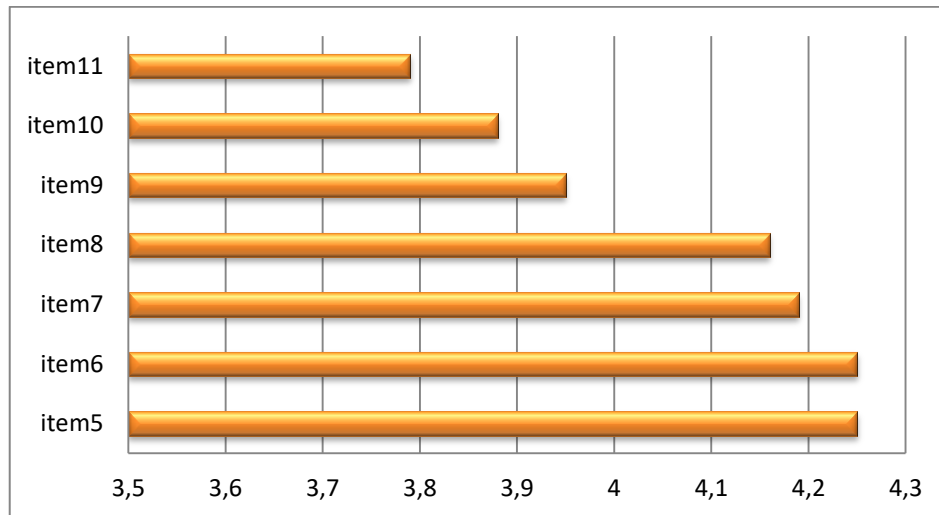
Thus, these data confirm the students' positive attitudes towards the integration of the reading and writing skills into the oral expression session (item6). They also think that the second semester oral expression session is more informative and beneficial (item3), and helped them develop their communication skills (item 11). As for the assessment practices, it seems that the students developed positive attitudes towards the IPA assessment model (item 2).

Expectedly, and since the students hold strong positive attitudes toward the oral expression second semester sessions contents, activities, feedback and assessment, they showed a very negative attitude towards the old teaching method they received formerly in the first semester (item 4). Since item 4 is a negative

item, the mean value $\bar{X} = 3,79$ is reversed to belong to high disagreement interval which, in turn reflects a negative attitude. The means values and levels of attitudes among the students are clearly displayed in graph 30.

Graph 25.

Students Attitudes towards Oral Expression Session graph



Clearly, all the items mean values are high ranging from $\bar{X} = 3,5$, as a lowest boundary to, $\bar{X} = 4,3$ as a highest value, which are both values belong to the intervals of agreement that, in turn represent the intervals of positive attitudes.

4.5.2 Students' Attitudes towards the IPA Activities and Outcomes

The second set of data represented in table 46 are related to the students' attitudes toward the activities used in the IPA model, as well as the relevant outcomes. Again, the items were ordered from the highest mean value to the lowest mean value and analysed in terms of the weighted mean values intervals.

Table 38.*Students' attitudes towards IPA activities*

Item	Statement	Mean	Std.D	Order
12	The activities we had in class helped me develop all my language skills (reading, writing, speaking and listening)	4,89	0,87	1
14	Reading and listening about the topic helps me to develop my knowledge about the topic I am going to present orally	4,87	0,82	2
15	Classroom discussions and comments helps me to adjust my ideas and correct my mistakes	4,60	0,62	3
19	My teacher feedback helps me to enhance my performance in all the activities	4,32	0,78	4
18	After reading, writing, and listening about the theme I can easily express about it orally	4,39	0,56	5
16	Debate and discussions help me develop my ideas and opinion about the topic	4,35	0,78	6
17	In the last phase of the session I can rigorously express my final opinion and thoughts about the topic	4,25	0,59	7
13	The first activities in oral expression session help to enhance my listening and reading comprehension	3,48	0,66	8

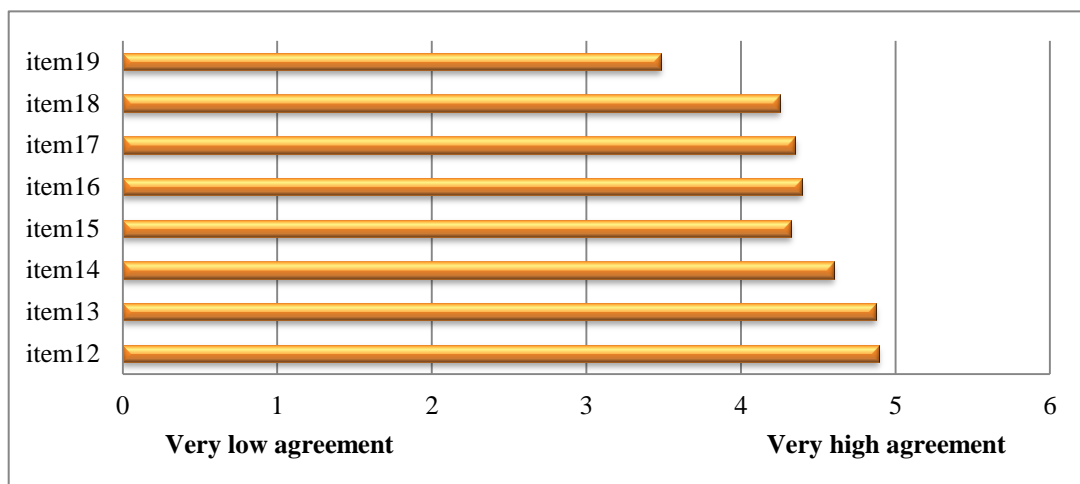
The analysis of the mean values items 12, 14, 15, 19, 18, 16 and 17 from $\bar{X}=4,89$, as the highest value, to $\bar{X}=4,25$, as the lowest value, confirms that they all belong to the higher agreement interval. In addition, the values of the standard deviation ranging from 0,87 to 0,56 points to the narrow spread of the students' responses from the mean, thus confirming that the majority of the students hold the same levels of agreement. Accordingly, these findings confirm the highly positive attitudes toward the activities of IPA. With regard to this, all the students highly agreed that the oral expression session helps them develop all language skills (item 12).

All the students strongly agreed on the fact that the use of the reading, listening, and writing activities helped them improve their oral expression skills (item 14 and item 18). They also highly agree that both teacher and peers' feedback through class discussions and debates helped them to improve their performance (item 15 and item 19). The interpretive stage of the IPA model was positively considered by the participants (item16). The participants hold strong positive attitudes toward the representational stage in which they orally present their task (item 17).

Last but not least, item 13 mean $\bar{X}=3.48$, which is a value that belongs to the second agreement interval, yet it also indicates positive attitudes. The Std. D value (0.66), being a very small value, confirms that most scores are around the mean value. Thus, the students hold a positive attitude towards the first stage of the IPA, confirming that all the stages are positively assumed by the students. Graph 26 clearly exhibits items mean values.

Graph 26.

Students' Attitudes towards Assessment IPA Activities and Outcomes graph



Overall, the findings from the students' attitudes questionnaire uncovered the unquestionable high levels of attitudes towards the used IPA as an assessment model in the oral expression session.

4.6 Analysis of students IPA performance

As part of examining the feasibility and usability of the IPA in assessing students' proficiency, the researcher used the IPA data as indicators of the levels of performance of the experimental group. The results obtained from the IPA were compared to the proficiency test data to examine the credibility of the IPA results. The following table summarised students' four skills performance as provided by the IPA.

Table 39.

Students' performance in the IPA Descriptives

	N	Min	Max	Mean	Std. Deviation
Reading	26	8,00	15,00	14,34	2,25
Writing	26	7,00	14,00	14,33	2,41
Speaking	26	9,00	17,00	14,36	1,89
Listening	26	10,00	16,00	14,34	2,03
Valid N	26				

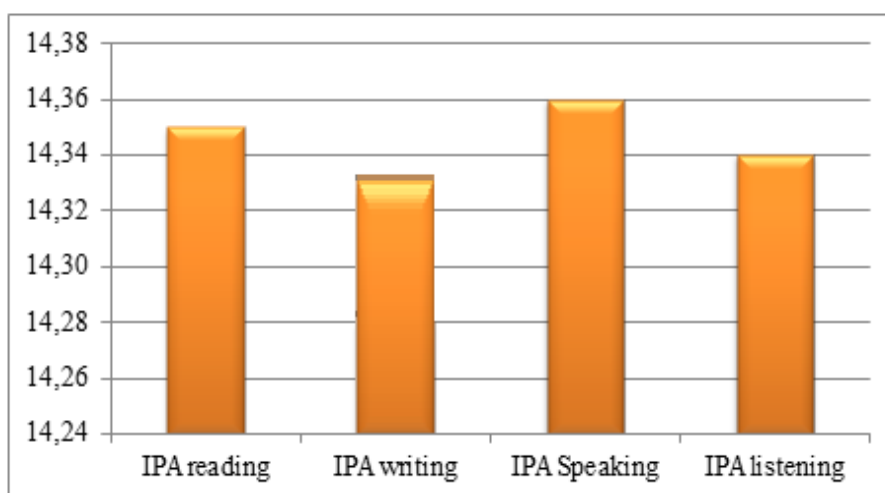
From table 47, which examines the mean values of the four skills, it seems that the students performed well in all the skills. With a range of eight between the highest score obtained by students in speaking (Max=17), and the lowest score (Min=9), and a mean of $\bar{X}=14,36$, revealing a good performance among the students, and a Std.D=1,89, revealing

close performance to the overall level, it can be said that the students' performance in speaking was a good performance.

Similarly, with a range of six between the highest score obtained by students in speaking (Max=16), and the lowest score (Min=10), and a mean of $\bar{X}=14,34$, revealing a good performance among the students, and a Std.D=2,03, revealing close performance to the overall level, it can be said that the students' performance in listening was a good performance.

Graph 27.

Students' performance in the IPA



The graph display confirmed the finding and reveals similar performance in the four skills is nearly similar to the results obtained from the proficiency test in the post testing phase. This corroborates that the evaluation provided by the IPA is valid.

4.7 Validity of the IPA

Validity of the IPA evaluation of the experimental group was examined through the use of correlation analysis. Exactly, concurrent validity is used through the running

correlation analysis of students' performance in the IPA and their performance in the proficiency post-test. The statistical analysis in this phase of the study was guided by one hypothesis:

- There is no statistically significant correlation between IPA score and the proficiency test post test scores.

The first correlation analysis was between total IPA score and the proficiency test post test scores. Table 48 represents the value of the correlation coefficient.

Table 40.

Concurrent validity of the IPA test (total scores)

		Post-test proficiency test	IPA
IPA	Pearson Correlation	,810**	1
	Sig. (2-tailed)	,000	
	N	26	26

** . Correlation is significant at the 0.01 level (2-tailed).

Apparently, there exists a statistically significant strong correlation between students' performance in the proficiency post-test and the IPA performance at the level $p=0.01$. This is confirmed by the value of $r=0.81$ that is significant at the level of $p=0.01$, which is a high value confirming a strong positive correlation. This finding confirms the validity of the IPA evaluation.

A further correlation analysis was conducted between the four skills performance in both tests. Table 49 represents the values of r in this second analysis.

Table 41.*Concurrent validity of the IPA test (the four skills)*

		Reading (PT)	Writing (PT)	Listening (PT)	Speaking (PT)
Reading (IPA)	Pearson Correlation	,773**			
	Sig. (2-tailed)	,002			
	N	24			
Writing (IPA)	Pearson Correlation		,667**		
	Sig. (2-tailed)		,040		
	N		24		
Listening (IPA)	Pearson Correlation			,580**	
	Sig. (2-tailed)			,002	
	N			24	
Speaking (IPA)	Pearson Correlation				,676**
	Sig. (2-tailed)				,000
	N				24

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.01 level (2-tailed).

Interestingly, there exists a statistically significant strong positive correlation between students' performances in both tests. From the examination of r values, table 49 above, the values of validity coefficient are statistically significant at the level $p=0.01$. It is also noticed that the values range from 0,850, as a lowest value, to 0,773 as a highest value. These values represent high values of validity. Accordingly, we can say that the evaluation of the students' performance using the IPA model of evaluation is valid.

4.8 Synthesis of the Findings

Throughout this chapter, various aspects of the data collection phase were presented, then the obtained data was analysed in detail. First, a discussion of the sample selection was presented, that includes the normality distribution and homogeneity of the research sample using the Kolmogorov-Smirnov and Shapiro-Wilk normality tests. After that, and since the test validity and reliability are two unavoidable qualities of measurement, the researcher presented in this chapter, the steps undertaken to ensure the IPA model validity and reliability during the piloting phase.

Moreover, and for the analysis of the obtained data, this chapter started with the analysis of the teachers' questionnaire which covered their perceptions of and practices about the oral expression teaching and assessment. The first part of the analysis tackled teachers' perceptions of and practices about teaching oral expression, and the second part probed their perceptions of and assessment practices vis-à-vis oral expression module. Then, throughout this chapter, the researcher provided a detailed analysis of all the data obtained for the present study to test the hypotheses put forward.

During the experimental phase, data were gathered from the pre- and post-testing stages and analysed through the use of the T-test for independent samples and T-test for paired samples to test the study's hypotheses. The first step involved testing the null hypothesis related to the pre-testing phase which suggested that there is no significant difference between the control and experimental groups' performance in the pre-proficiency test. The findings

revealed that there was no significant difference in the performance of both groups, thus leading to the acceptance of the null hypothesis.

Moreover, the third Hypothesis, which stated that: There is no statistically significant difference between the experimental group performances in the pre and post-test, was tested through the use of Paired Samples T-test to compare between the experimental group pre- and post-test's performance. Data obtained in this analysis revealed a clear improvement in the experimental group performance after the integration of the IPA. Therefore, this hypothesis was rejected.

The last difference analysis was between the control group performance in the pre- and post-test, without being subject to treatment, was devoted to test the fourth null hypothesis stating: There is no statistically significant difference between control group performances in the pre- and post-test. This hypothesis was confirmed through the use of paired- sample t-test since the results revealed that there is no significant improvement in the proficiency levels among the control group and their levels were almost the similar. This analysis was necessary to ensure the internal validity of the experiment.

Thus, from these findings, the researcher concluded in this chapter that the main hypothesis is accepted confirming that the implementation of IPA assessment model contributed to the promotion of language skills' proficiency levels among the participants.

The final phase, in the research process of the present study, was dedicated to test the washback of the implemented IPA assessment model on the students' attitudes, and motivation towards teaching, learning, and assessment in oral

expression through the use of a questionnaire administered to the experimental group. The data obtained from this students' questionnaire was divided into two sections. The first section detected students' attitudes towards the oral expression session after the integration of the IPA. The second section sought to understand students' attitudes towards the IPA activities and outcomes.

Ultimately, as a part of examining the feasibility and usability of the IPA in assessing the students' proficiency, the researcher used the IPA data as indicators of the levels of performance of the experimental group. The results obtained from IPA were compared to the proficiency test data to examine the credibility of IPA results. The findings revealed similar performance in the four skills is nearly similar to the results obtained from the proficiency test in the post-testing phase. This confirms that the evaluation provided by the IPA is valid. Also, the validity of the IPA evaluation of experimental group was examined through the use of correlation analysis. Exactly, concurrent validity is used through running correlation analysis of students' performance in the IPA and their performance in the proficiency post-test. The first correlation analysis was between total IPA score and post test scores. values obtained in this analysis represented high values of validity. Accordingly, it was concluded that the evaluation of students' performance using IPA model of evaluation is valid.

Conclusion

This chapter presented the data analysis, interpretations and synthesis of the findings obtained from present study. It started up by testing the “Normality distribution and Homogeneity” of the research sample, then, testing the IPA proficiency test reliability and validity.

After that, the researcher presented a detailed description and analysis of the data obtained from the “Teachers Questionnaire”. The analysis provided valuable insights into their perceptions of and practices in oral expression module. This step was followed by analysing teachers' perceptions and practices in the oral expression assessment, which lead to understand the factors that teachers consider when assessing student performance in this module and revealed their readiness for change into alternative assessment.

Moreover, this chapter included the analysis of the data obtained from the experimental phase which revealed that the implementation of the IPA had a positive effect on enhancing learners' language skills proficiency. This finding is consistent with previous research that has demonstrated the effectiveness of the IPA as an assessment tool for measuring language proficiency.

Furthermore, the analysis of data obtained from “Students Questionnaire” related to students' attitudes towards oral expression's new sessions, which included the implementation of the IPA, provided important information about the effectiveness of this assessment model in promoting students' engagement and motivation.

Finally, the feasibility, usability and validity of the IPA in assessing students' proficiency were thoroughly examined then confirmed through

correlation analysis that included the analysis of students' IPA performance in regards to the IELTS results.

Overall, the findings of this study provide valuable insights into the effectiveness of the IPA as both a pedagogical and assessment tool for assessing and enhancing students' proficiency in oral expression. These findings have important implications for teachers and researchers seeking to improve language education and promote effective teaching practices.

A further discussion and development of these results will be presented in the next chapter within a framework of related literature. It will also cover the pedagogical implications, limitations, challenges, and suggestions for further research.

CHAPTER FIVE:

Discussion of the Findings, Pedagogical Implications, Limitations, and
Further Research Recommendations

Introduction

This chapter is devoted to the discussion of the overall findings obtained from the different tools and stages of this study. The findings are presented and discussed within the framework of the related literature. The discussion interconnects all the obtained findings so as to offer clear and direct answers to the study questions. In doing so, this chapter confirms the effectiveness of implementing the IPA model to assess and enhance EFL learners' language proficiency and in boosting their motivation in the oral expression session. Additionally, this chapter sheds light on the study outcomes' and pedagogical implications, then presents all the limitations encountered by the researcher. Finally, this chapter suggests a number of recommendations for further research.

5.1 Teachers' Perception of the Integration of the four skills in Language Assessment

The findings related to the first research question, which investigates the current assessment practices among Algerian EFL teachers in oral expression module, sheds light their perceptions about the oral expression module teaching and their current assessment practices.

5.1.1 Teachers' Perceptions of the Oral Expression Module Teaching

Since the aim behind integrating the IPA model in the oral expression session is to enhance students' proficiency levels, including all four skills (reading, writing, listening, and speaking), it was necessary to investigate the oral expression teachers' beliefs about integrating the four skills in this course.

Findings of the first item revealed that all participant teachers agree that teaching oral expression course aims exclusively at developing students' speaking and listening comprehension skills. Building on this perception regarding the objectives of teaching the oral expression module, teachers held the perception that the focus is only on two skills and that there is no need to integrate all four skills in teaching this course. This finding is confirmed through the responses to the next item in which teachers ranked "speaking and listening" as the main skills to gain higher focus while "reading and writing" were ranked as the least needed, and the less important skills, in the oral expression session. They indicated that they do not usually integrate "Reading and Writing" in this module. Only a few of them explained that they use them as helping activities to prepare students for the speaking activities, or as homework to prepare for the coming sessions.

However, to develop learners' listening and speaking skills, other skills including; reading and writing are needed. According to Oxford (1990), the acquisition of a new language inevitably involves developing and linking the four language skills in varied degrees. Likewise, Brown (2000) stated that "the richness of integrated-skill courses gives EFL students greater motivation that converts to better retention of principles of effective speaking, listening, reading, and writing" (p. 233). He added that, productive and receptive skills are "two sides of the same coin" where each one of the skills helps in reinforcing the other, thus, neglecting the link between them is neglecting the richness of the language (p.234). On the same vein, Peregoy and Boyle (2001, as cited in Chen, 2007) confirmed that the integrated use of language skills in subsequent use, in classrooms, helps learners to promote their own learning and they become able to

use the target language for functional and meaningful objectives, which will certainly improve their language proficiency by time.

Several other studies indicated that each of the skills is needed in the process of developing the other and asserted that they have to be integrated altogether throughout the language learning process as a key element to success because they are inseparable and interdependent (Oxford, 1990; Tierney & Shanahan, 1991; Alderson, 2000; Brown, 2000; Chen, 2007).

For the use of reading in the oral expression course in order to develop the speaking skills, Grabe (1991) pointed out that, in language learning, developing speaking skill is strongly related to reading that supplies learners with sufficient vocabulary and structural knowledge (p.379). Thus, starting an oral expression course with a reading task offers learners an opportunity to acquire new facts, knowledge, contexts, and grammatical structures, as well as to enlarge their vocabulary. This will help them later to speak fluently and express their ideas effectively (Hamrayevna et al., 2019). Similarly, Krashen and Terrell (1989) argued that reading also empowers language learners' understanding and listening comprehension. It plays a crucial role in improving language competence (p.131). Therefore, teachers must bear in mind that one of the most useful ways to improve students' communication skills is reading.

The relationship between the speaking skill development and the writing skill is similarly highlighted. Preparing students to speak sometimes requires asking them to prepare a written draft, otherwise, they would find difficulty in speaking fluently. Therefore, incorporating writing tasks in the oral expression

session before speaking tasks or presenting oral works can be an effective way to prepare students to speak (Dupont et al., 2020).

However, the teaching practices in the Oral Expression course in Algerian universities is limited to two skills “ Speaking and Listening” being the objective and the sued skills in teaching this module. This was highlited in the first items confirming that developing them is the targeted objective of this course. Through the reponses of teachers a clear neglection of the importance of integrating the other two skills to support and develop learners’ Listening and speaking was notived. Hence, the IPA can serve as both a pedagogic and assessment tool that serves in integrating the four skills in a seamless manner throughout its three stages in the targeted subject. According to Madison (2019) each of the integrated skills help in developing the other in a cyclical approach and all serve in the development of the targeted skills. In this study, the selected subject is the oral expression, thus, the integrated skill would definitely serve in the development of the targeted skills which are the speaking and listening skills.

5.1.2 Teachers' Perceptions of the Oral Expression Activities

For the activities involved in the oral expression course, the questionnaire choices were built upon Harmers' (2001) suggestions of the activities that could be implemented in oral expression classes, including: Listening comprehension activities, role play, classroom discussions, interviews, debates, communication games, dramatisation, dialogues, and reading time.

Findings revealed that the listening comprehension activities are the most used activities in the oral expression course. In fact, listening is crucial for developing language proficiency. Most of the previous research, which dealt with

developing the speaking skill of foreign language learners, indicated that, speaking and listening are strongly related. According to Krashen's model (1985) "the input hypothesis", the effective acquisition of a language is based on the "comprehensive input" introduced to the learner i.e. the sufficient, comprehensible, relevant, and contextualized input presented to the learner in the forms of listening material. Respectively, Rivers (1981) specified that "Speaking does not of itself constitute communication unless what is said is comprehended by another person. Teaching the comprehension of the spoken speech is therefore of primary importance if the communication aim is to be reached" (p. 22). Therefore, listening comprehension activities are an essential part of the oral expression course. In the present study, the listening comprehension phase comes as the first part of the IPA, towards assessing and improving learners' proficiency.

Additionally, speaking activities, including discussions, question-answer participation, dialogues, storytelling, and interacting with various topics, are also widely used by the respondents. Equally, movies and songs were used. Suban (2021), as a conclusion to his study on using speaking activities to teach and promote learners' speaking skills in EFL classrooms, affirmed that "Optimising classroom time with appropriate learning activities can be a better choice. Practical activities such as dialogue, surveys, picture description, picture differences, role plays, story completion, and storytelling can be fun and formative, and can help to improve language learning, in general, and to promote learners' speaking skill, specifically" (p.49). Accordingly, the second interpersonal phase of the IPA model, of this study, provided an opportunity for students to speak while they discuss, interact with each other's views, and present their ideas.

Furthermore, the use of Role play, in the findings among respondents, was moderate. Various previous studies reflected the importance of this activity in EFL classrooms as they offer significant opportunities to students to use the language to perform and communicate in real-life contexts (Murdibjono,2001; Danasaputra, 2003; Diani,2005; Ur, 2012; Larsen-Freeman et al., 2013). In the IPA model, implemented in this study, role play activity was among the selected activities in the final presentational phase.

However, according to the findings, group work was not widely used in spite of the importance of this activity for the oral expression course. With the use of group work and pair work, students are likely to talk and interact. Successful and effective group work can offer great opportunities for students to participate in open conversations, discuss different ideas, convey and comprehend meaning, as well as to develop social skills such as taking turns in conversations. (Larsen-Freeman et al., 2013). Therefore, for effective use of classroom time, language teachers should engage their learners in pair work and group work (Ur, 2012). Other activities that were proposed by the participants include: debates which are a part of the speaking activities, oral presentations, and reading circles. All the above mentioned activities are among the various tasks of the IPA which can be involved as both pedagogical practices and formative assessment tools (Adair-Hauch et al., 2016)

Interestingly, the majority of the participants confirmed that they make sure that the selection of the topics and activities for the oral expression session suits their learners' interests by giving them a chance to choose them. This approach, to the selection of themes, is significant in creating a motivational atmosphere in

the language classroom. It helps in providing energetic and engaging activities that allow learners to use the language comfortably to express their opinions and to participate (Hungyo et al., 2009). It is conventionally identified that the level of interest among learners promotes students' motivation and increases their abilities to comprehend and interpret (Dristas et al., 1995). Therefore, the IPA as a pedagogic and assessment model, according to the manual (2016), "provides opportunities for students to select their own authentic texts of interest and demonstrate their comprehension and interpretation of them" (p.93).

5.1.3 Challenges to achieve proficiency in Language skills

Among the encountered hindrance, according to the participants, is the students' weakness in language proficiency, which constitutes the most problematic aspect. Other aspects stated by teachers include:

- the lack of motivation among students;
- overcrowded classes;
- difficulty in selecting types of teaching activities that help in motivating students;
- the lack of logistics including labs and ICTs tools;

For the last point in the teacher responses, the integration of the teaching resources including technology and ICT within the EFL classroom holds the potential to foster a dynamic learning environment that simulates authentic social interaction. The use of these technologies provides an opportunity to engage language learners in a more interactive and stimulating manner. In this regard, Richards and Renandya (1990) claimed that teaching materials play a crucial role

in EFL classrooms. The quality of these teaching materials can strongly affect the quality of the language learning process as well as students' motivation. Correspondingly, language teachers have to carefully select the teaching materials integrated into their classrooms.

The use of technology and ICT's plays a key element throughout the three phases of the IPA. During the interpersonal phase, for instance, students should be exposed to authentic audios or videos where they should listen and interact to authentic materials. Then, while presenting their final works in the presentational stage, students are required to bring videos, power point (PPT) presentations or any products they created for the session using audio-visual materials. In the present study, this use of these authentic materials including videos, PPT presentations and pictures had significant effect on learners' motivation and participation. It was an effective method to attract their interests and attention to the session.

5.1.4 The Current Assessment Practices among Algerian EFL Teachers in the Oral Expression Course

The second part of the results is about the assessment practices in the oral expression session. Participants specified that oral presentations are the most used assessment methods in this course. Other methods include listening comprehension activities in the form of paper-pen tests, classroom interactions, and the least used is classroom discussion whereas both reading comprehension and written expression received no agreement on the possibility of using them to assess students in the oral expression module. As to the alternative assessment

methods in oral expression, only few teachers confirmed the use of methods including performance-based assessment such as group projects and role plays.

For feedback, as an important part of the assessment process, teachers showed that they always provide feedback to their students in all the sessions of oral expression. Only a few of them do occasionally provide feedback on their students' performances. According to the findings, the most used type of feedback is delivered to students in the form of oral notes whereby they provide direct correction of the mistake. This correction involves correcting: pronunciation, grammar, and speaking mistakes. However, only a few of the participants use peer feedback. In the IPA, feedback plays an important role in enhancing students' performance in each phase (the interpretive, interpersonal, and presentational) and every task (Sedor, 2022).

Additionally, the difficulties that teachers encounter when assessing oral expression include the following;

1. The large number of students per group;
2. The limited number of allotted sessions;
3. The weak levels of proficiency among the students.

As to the integration of the four skills in the process of evaluating students' overall performance throughout the oral expression sessions, most of the respondents confirmed that they never use the four skills in their assessment practices because they find no efficiency and no need for integrating them in the evaluation process. Thus, they confirmed that they evaluate only the speaking and listening skills. Yet, a few of them confirmed that they occasionally integrate the skills. This finding is along with the findings related to the objectives of teaching

oral expression that revealed that the objectives are the listening and speaking skills.

Most of respondents did not agree with the use of traditional methods of assessment as the main tools to assess their students' performance in the oral expression module (including paper pen tests). They think that the use of traditional methods of assessment is not sufficient.

In light of these findings, it seems that even though teachers lack awareness about the integration of the four skills in the teaching or the assessment in the oral expression skills. They are aware and open to the importance of integrating alternative methods of assessment. Thus, the suggested intervention based on the IPA can be an effective alternative assessment method that teachers can use in the oral expression as it holds not only the integration of the other skills but also effective feedback provision techniques, the use of ICTs, and methods of performance evaluation.

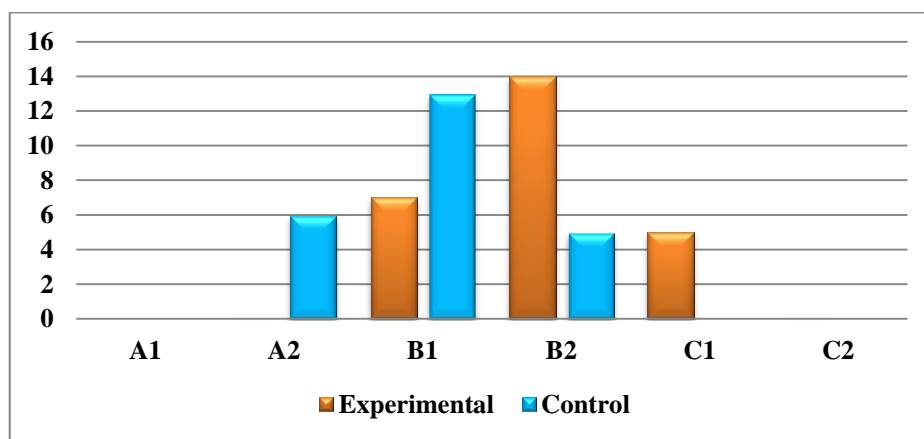
5.2 The Impact of IPA on improving students' language proficiency

First, findings from the post-testing stage revealed that there exists a statistically significant difference between the control and experiment groups' performance in the proficiency test. Similarly, a statistically significant difference between the experimental group's pre- and post-test performance was detected. These findings confirm the impact of using the IPA model in the oral expression session which served in improving students' proficiency in the four skills.

With a significant mean difference as indicated in table 22 (See chapter 4, section 4.2.1) that is clear from the mean values (experimental $\bar{X}=5.08$, Std.D=0.67 and control $\bar{X}=4.07$, Std=0.87), it is confirmed that experimental group performed better than the control group reflecting the possible positive impact of the IPA model use in the development of students' proficiency. The difference in performance is clearly displayed on graph 19.

Graph19.

Post-test Proficiency Levels Distribution among Experimental and Control Group



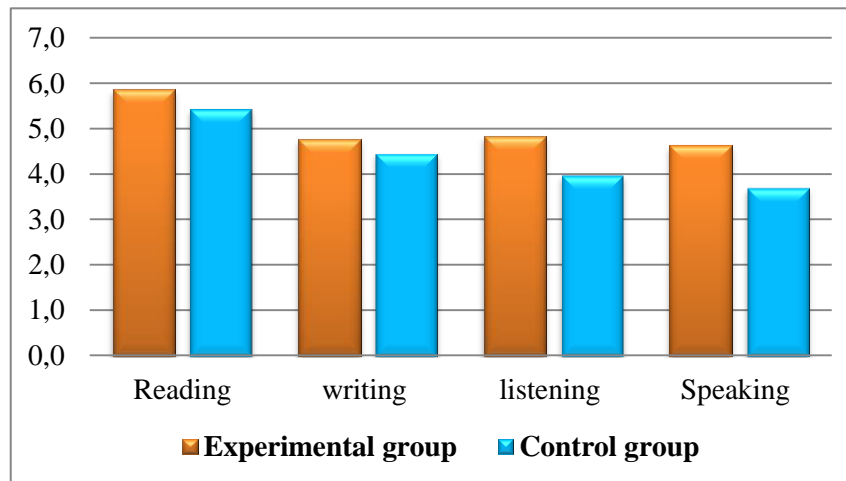
The clear difference between proficiency levels distributions is well displayed. The experimental group participants considerably exceeded the control group participants in the good level B2, developed to the advanced C1 level of proficiency and no one was categorized in the B1 level. Thus, marking a significant improvement in the proficiency level compared to the control group's performance.

To specifically understand the impact of the IPA on the proficiency levels, performance in the four skills targeted by the proficiency test for both

experimental and control groups were compared. Graph 20 (from chapter 4) exhibits the differences in the four skills performance in the post-test.

Graph20.

The differences in the four skills performance in the post-test.

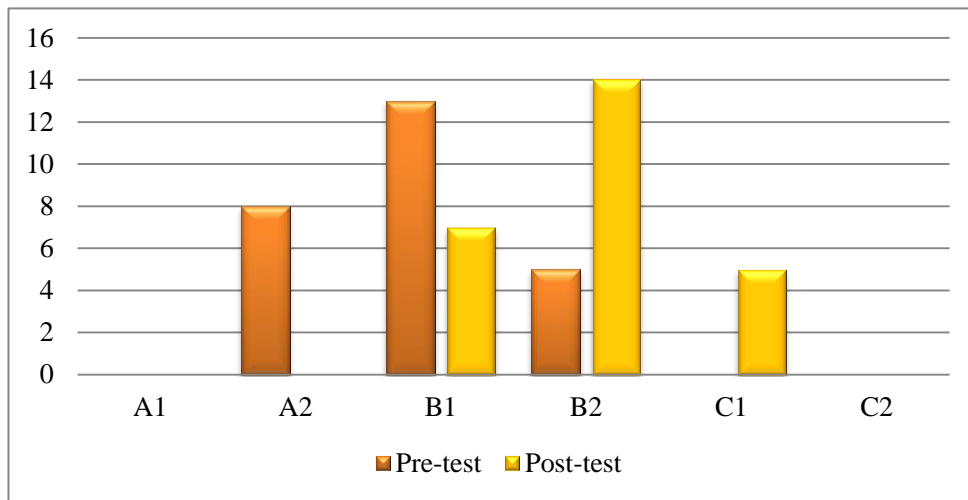


The detected difference is noted in the four skills showing that experimental group participants have notably improved their proficiency in the four skills. Remarkably, the improvement in listening and speaking skills, being targeted skills by the oral expression module, is highly significant compared to the reading and writing skills.

To confirm this finding, comparison of experimental group participants' performance in the pre- and post- proficiency tests was conducted revealing a significant improvement in their proficiency levels. From the mean values of both performances (pre-test $\bar{X}= 4.50$, Std=0.83 and post-test $\bar{X}=5.08$, Std=0.87), the improved performance is confirmed. Furthermore, as displayed on graph 22, levels of proficiency are significantly improved as A2 level are absent, B1 level had reduced, B2 level has remarkably increased and a new advanced level has been detected.

Graph28

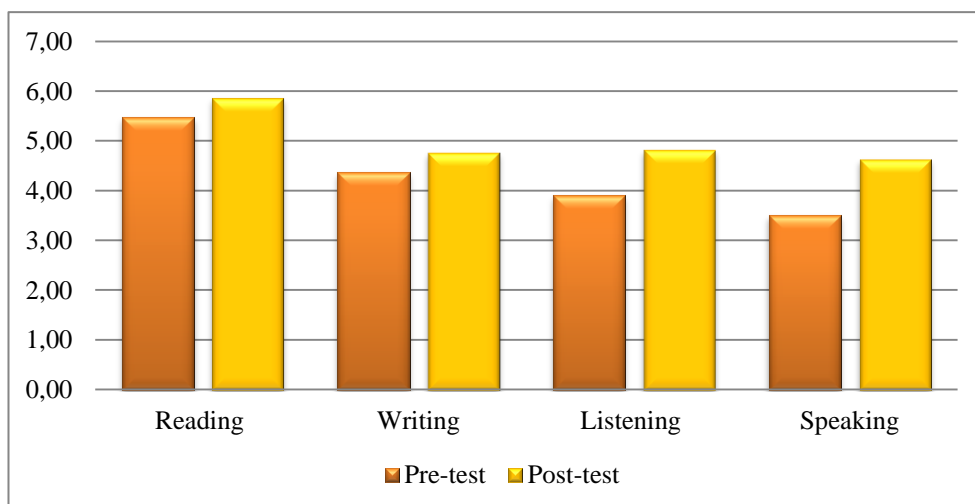
Differences in the pre-and post-test proficiency levels distribution



As to tracking the development in the four skills proficiency the following graph clearly represents the marked differences in the proficiency of the four skills that unquestionably confirm the positive impact of the use IP in the improvement of the four skills performance. Specifically, the most significant difference is detected in the listening and speaking skills.

Graph 29.

Differences in experimental group pre- and post-test performance



In light of the aforementioned summary of post-testing findings, the effectiveness of using the IPA strategies to improve students' language proficiency in the four skills is confirmed. Similarly, the study by Lee and Lim (2020) was conducted to investigate the effectiveness of using the IPA in improving the English proficiency of Korean learners across all four skills. The results obtained from their study displayed that the experimental group, who received instruction on the IPA, showed significant improvement in their English proficiency in all four skills compared to the control group. Additionally, they affirmed that the IPA provided their students with more opportunities to practice and develop their language skills in a more integrated and authentic way (Lee, et al., 2020). This positive impact, as explained by (Sedor, 2020), is the result of real-world communication activities and provided feedback.

Unlike this study's focus, on exploring the impact of IPA on the four skills, others studied targeted its impact on separate skills like the speaking skills. According to Kurniawan et al., (2022) study findings, implementing this IPA model can contribute to the improvement of students' speaking skills. This finding confirms, again, the effectiveness of using IPA in the oral expression session since speaking skills is viewed as its main targeted skill.

This study finding goes along with current trends in assessment practices, which calls for closer connection between teaching/learning and assessment practices, and the positive impact assessment should play in improving both (Adair-çHauk et al., 2016). Thus, and since the IPA use proved to have a positive impact on this study participants' levels of proficiency, it is assumed that IPA as an assessment method has a constructive positive washback effect on EFL

students' development of the four skills if integrated in the EFL language classroom, specifically in this study, the oral expression session. This confirms currently acknowledged novel roles of assessment practices in promoting learning and teaching (Adair-Hauck et al., 2013). This also confirms the washback effect on student learning and classroom instruction (Adair-Hauck et al., 2006; 2011). Further washback effect on students' attitudes is explained in section 4.

In explaining the role of IPA in improving teaching and learning, Adair-Hauck et al. (2013), asserted that the followed assessment protocol of the IPA model that encompasses real-life authentic documents and tasks, models of expected performance, and co-constructive feedback, helps to connect teaching, learning and assessment activities in one action targeting one objective. In the present study, for instance, with the use of authentic documents in the first phase (interpretive), the students were so interested to read a native English citizen describing the city of London and its landmarks. Firstly, because the theme was of their interest, and secondly, the target city described in the text was their choice. This authentic text, also, helped them to gain new knowledge and vocabulary regarding the selected topic that supported their speaking performance in the next phase. After that, during the interpersonal phase, the warm-up activity, which involved listening comprehension tasks using an authentic video, remarkably attracted their attention. Using this video as a model, the students were aware of their expected responses. Respectively, they participated and responded efficiently.

Additionally, constant co-constructive feedback presented in these two phases provided strong support to each student to figure out his/her weaknesses and enhance them, to spotlight upon language flaws for further improvement.

Feedback from the interpretive phase, helped students to better understand the authentic document, learn how to skim and scan the text for different types of information and ideas, and how to read between the lines. These reading strategies served as a support to enhance their reading skill. Additionally, feedback provided in the interpersonal phase, which was constructed with the students, included various aspects of the language such as: consistency, comfort, quantity and organization of language discourse, connecting sentences to extended discourse, interactivity, clarifying meaning in breakdowns in communication, comprehensibility, grammatical accuracy, pronunciation, appropriate vocabulary, and degree of fluency. This feedback loop of the interpersonal phase played an integral role in developing the students' speaking skill. Hence, at the final phase of the IPA (the presentational), the students were notably ready, well prepared, and not stressed from this evaluation, even though this phase was to be considered as their summative assessment, because they have already got an idea about the expected performance and their weaknesses that should be improved in the previous sessions. The results of their written and spoken presentations were impressive since most of the participants reflected a remarkable improvement.

Similarly, Sedor (2020) confirms the importance of feedback stating that the different types of feedback are crucial in the teaching\ learning process. They have strong effects on students' performance and skills. This effect can be either positive or negative. However, constructive feedback is not just advantageous and favourable for the learning process, but also supportive and motivating.

5.3 The Use of IPA to assess language proficiency

It is important to mention that the focus of the intervention was not only on testing the effectiveness of IPA in improving language skills proficiency, but also testing its potential use to assess proficiency in the four skills. As part of examining the feasibility and usability of the IPA in assessing students' proficiency, the experimental group students' performance was tested by both the proficiency test and the summative results of the IPA model.

The first step was to examine the levels obtained from the IPA. The descriptive statistics obtained from the IPA indicated similar performance descriptions to the performance of the proficiency test (IELTS), revealing that the students' performance was similar in both tests, thus confirming the validity of the IPA to assess their proficiency. The latter was further examined through the use of concurrent validity by running a correlation analysis of students' performance in the IPA and their performance in the proficiency post-test. The first correlation analysis between total IPA score and post-test confirmed the existence of a strong positive statistically significant correlation ($r=0.81$ at the level $\alpha<0.05$) between the performances of both tests confirming the higher validity of the IPA summative procedures results.

Further correlation analysis was conducted between the four skills' performance in both tests. The values of r , ranging from 0,850, as the lowest value, to 0,773 as the highest value at the level of significance $p=0.01$, confirm the existence of a statistically significant strong positive correlation between students' performances in both tests. Accordingly, this strong correlation is another confirmation of the IPA validity in testing language performance. Thus,

along with the findings of related research, this study's findings confirm the validity of the IPA in testing language proficiency through the four skills through the use concurrent validity.

In line with the research on IPA feasibility and validity in testing language skills, it has proved its effectiveness by broadening the scope of assessment from testing single skills into considering the relationships among language skills which occur normally in the course real-world communication (Adair-Hauck et al., 2006).

Moreover, confirmed by Adair-Hauck et al., in their various studies and works on the IPA (2006; 2011; 2013, 2016), that designers of this assessment model have structured the IPA and provided: detailed explanations about the interpretive, interpersonal, and presentational tasks, clear rubrics which guide the task completion and scoring, and a full description of required feedback of this assessment. They prepared the IPA to meet the qualifications of a valid and reliable assessments tool to assess proficiency. The latter, as defined by Adair-Hauck et al. (2011) as: "the level at which students comprehend and interpret authentic texts in the foreign language, interact with others in the target language in oral and written form, and present oral and written messages to audiences of listeners and readers" (p.606). Keeping with IPA validity, Glisan et al. (2003) research proved that the IPA is a valid and reliable tool to assess proficiency in the three modes of communication with young Spanish language learners. It was also confirmed that the IPA can be a productive PBA tool to assess not only learners' learning but also teachers' practices (Davin et al., 2011).

Interestingly, another type of validity is also addressed in this study which is consequential validity which relates to the washback of IPA on learning and teaching. According to Morrow (1986), a test's validity should encompass the degree to which the

test has a positive influence on teaching and learning. Therefore, since this study overarching objective was to study the impact of IPA on proficiency, it has contributed to the study of the consequential validity of the IPA. Then, and since the findings confirmed its impact in improving students' performance and attitudes, it is unquestionably confirmed that the IPA model is a valid assessment measure. Similarly, according to Adair-Hauck et al., (2013), the goal of devoted research on the effectiveness of the IPA was to investigate the washback effect or consequential validity of the IPA on teaching and learning.

5.4 The Washback Effect of the IPA on Students' Attitudes

In addition to studying the washback of IPA on improving language proficiency, this study opts to examine further washback on students' attitudes towards the instruction and the oral expression session being the module through which the IPA was integrated through the use of a questionnaire to experimental group students.

5.4.1 Students' attitudes towards the oral expression session after the integration of the IPA

The finding from the first section revealed that the students had positive attitudes towards the oral expression session in the second semester compared to the first semester. All the students highly agreed that the second-semester's oral expression sessions helped them to improve their knowledge and their vocabulary, as was explained in the previous section.

Furthermore, the students highly confirmed that the oral expression session is more enjoyable and motivating. First of all, giving them the chance to select the theme and the topic they prefer was one of the factors that increased

their interest and foster their motivation for the session. Similarly, Sedor (2022) in his study, stated that by giving a chance to the students to modify their IPA task according to their preference, they were provided a feeling of comfort and motivation that led them to fluidly speak during the interactive tasks.

Additionally, since the students were EFL learners and very interested in learning the language, as well as the culture of English native speakers, using authentic videos and texts was very interesting and helpful for them. Another motivating element, in the second semester's oral expression sessions, is that each student is evaluated and provided with sufficient instant feedback to improve in the next session. Therefore, the resulted feeling of continuous improvement among the students motivated them. Moreover, the variety of tasks, involved throughout the three phases, has reduced the usual boredom felt during the previous oral expression sessions because of the traditional regular listening comprehension and speaking activities. This variety in tasks has also provided each student with an opportunity to speak, interact and express opinions. Thus these findings confirm the positive washback on students' motivation to improve their language skills. In the same regard, Glisan et al. (2007) and Davin et al. (2011) reported students' claim that the IPA tasks, modeling and feedback motivated them more to participate and engage in speaking tasks.

Moreover, as has been mentioned earlier, the students hold highly positive attitudes towards the teacher's feedback and the peers' feedback. They all agreed that both the teacher's and their peers' feedback, which was provided after class discussions and speaking tasks, helped them to improve their performance and skills. Unlike evaluative and judgmental feedback, the IPA takes a descriptive

approach to feedback that helps learners to understand what they need, spot their weaknesses, and flaws then help them to improve their performance (Adair-Hauck et al., 2003; Adair-Hauck et al., 2013; Zapata, 2016; Ching-Hsuan, 2018).

In the seminal study of Adair-Hauck et al. (2013), which is entitled “A Descriptive and Co-Constructive Approach to Integrated Performance Assessment Feedback”, the researchers highlight the importance of feedback in the IPA model, as well as to describe the characteristics of effective feedback. This feedback is also important in that it can be beneficial for all the students to learn from each other's mistakes, and then find out how to improve them. This can be done while listening and helping in the process of constructing feedback for each student in the class (Davin et al., 2011). In the study of Davin, et al., (2011), after IPA was implemented in fourth and fifth-grade classes with Spanish classes, it was found that learners held positive attitudes toward the use of descriptive feedback as well as collaborative work on the IPA tasks. According to the researchers, this was an indication of their understanding of the connection between their learning and assessment.

Furthermore, the results from Sedor's study (2022), which investigated how instructor feedback affects student performance on IPA assessments, revealed that feedback provided on the IPA helped in improving students' performance as it assisted them in recognising and spotting their weaknesses that should be improved and their strengths as well. His finding indicated that whenever the feedback was absent in the three phases, students' performance declined or remains constant. Therefore, he concluded that the feedback from the IPA

contributed to the students' success and it was significantly beneficial for them in the interpersonal and presentational tasks (Sedor, 2022).

Modeling, as a crucial element of the IPA (Adair Hauck & Troyan, 2013), was desirable and substantial for students since it helped them to have a clear image of their expected work and performance in this study. It was in the form of a video during the interpersonal task that served as both, an authentic material for the listening comprehension task, and a model of the expected performance for the presentational phase. Similar results from Sedor's study (2022) confirmed the importance of the models of the expected performance. He assumed that modeling is significant to clarify the expectations of the task and avoid students' misperceptions.

Adair-Hauck and Troyan (2013) described the process of modeling in their study, which was presented in their IPA implementation. Their finding depicted that the use of modeling helped their students to effectively understand the interpersonal task and its rubric. They also claimed that throughout the speaking tasks, modeling helped their students to adjust their performance and improve.

5.4.2 Students' Attitudes toward the Integration of the IPA as an Assessment tool

New attitudes towards assessment methods used in the oral expression module in the second semester, after using the IPA, were also explored. The findings confirmed students' positive attitudes towards the integration of the reading and the writing skills into the oral expression session. This attitude was depicted through their strong agreement on the fact that the use of the reading, listening, and writing activities helped them improve their oral expression skills.

They also think that the integration of all the skills in the second semester oral expression session made it more informative and beneficial and helped them develop their communication skills as well as their knowledge repertoire.

In each phase of the three modes of communication, the IPA makes use of one of the language skills (Reading, Writing, Listening, or Speaking) to support the other skill in the next phase. Thus, each of the integrated skills helps in developing the other in a cyclical approach. Reading, for instance, at the first interpretive stage helped to collect some accurate language structure, vocabulary, and knowledge. Then, listening to authentic language in a video, during the second interpersonal phase helped to support their pronunciation, grammatical accuracy, appropriate vocabulary, degree of fluency and strategies for clarifying meaning. After that, students were asked to prepare a written work for the presentational phase before presenting. This writing task assisted them to prepare accurate work, plan and organize their language, select the appropriate language, and support their speaking ability during their presentations. Hence, in the presentational phase, students have remarkably improved in their speaking skills.

Accordingly, all the acquired repertoire of knowledge adopted through the receptive skills (reading and listening) plays a vital role in enhancing the students' productive skills (speaking and writing) since it allows students to use what they learned to participate and speak accurately in the interpersonal phase then write, create and present in the presentational phase (Madison, 2019).

Similarly, in the study of Adair-Hauck et al., (2013), the students noted that the different speaking tasks used in the IPA helped them to adopt a wide range of vocabulary, then supported them to speak and interact with questions and

discussion fluently (Adair-Hauck et al., 2013). Additionally, according to their study, learners' performance in the interpersonal mode was not as good as that of the presentational. In the latter learners perform and presented better; however, in both modes learners have met the expectations (Adair-Hauck et al., 2013).

Correspondingly, the present study revealed almost similar results. Where the students performed better in the second and third phases during the speaking and presenting tasks, however, in the interpretive reading tasks their performance was not as good as the other phases and lower results were marked.

Likewise, the study of Davin et al., (2011), lower levels of performance in the interpretive mode among the participants were marked, however, they performed best in the presentational mode on which fluidity and advance were noticed (Davin et al., 2011).

As for the assessment practices, it is highly agreed that positive attitudes and perceptions of assessment can affect either positively or negatively on students' learning process and academic success. Positive perceptions, for instance, is a key element for motivation, as it was explained earlier in this section. Students with negative perceptions toward assessment may suffer from stress, anxiety and lack of confidence and motivation during the assessment process. (Entwistle & Tait, 1990; Struyven et al., 2005; Syaifuddin, 2019).

Findings of this study showed that the students developed positive attitudes towards the IPA model and its activities. It was confirmed that all phases of the IPA interpretive, interpersonal, and presentational were positively assumed by the students. In the same vein, various studies conducted to probe the perceptions and

attitudes on the implementation of the IPA, as an assessment model, among both students and teachers, were mainly positive (Sedor, 2022; Altstaedter and Krosch 2018; Martel 2018, Zapata, 2016). Zapata's study (2016), for instance, regarding the washback of IPA on learners' attitudes, revealed that learners integrated in the IPA classes liked the assessment tasks as they were able to link classroom instruction and assessment. Therefore, they reflected highly positive attitudes towards the IPA.

Both teachers and students' perceptions of the implementation of the IPA were explored in the related literature. Among these studies are Glisan et al. (2007), at the post-secondary level, and Davin et al. (2011), in an early foreign language program. Both studies involved addressing a survey research to investigate the students' perception after implementing the IPA in their classrooms. Findings of both studies revealed highly positive attitudes and perceptions.

Expectedly, and since students held strong positive attitudes towards oral expression in the second-semester sessions' contents, activities, feedback, and assessment, they showed a negative attitude towards the old teaching method they received formerly in the first semester.

5.5 Pedagogical Implications

The assessment process is a significant element for both learning and teaching since it is the backbone of these processes. Thus, understanding the assessment is crucial for an effective selection, and effective use of its various types, tools, and methods. For this reason, this study offers a holistic framework

explaining the assessment process in the EFL classroom through a demonstration of its components, stages, and different methods and tools. Then, it focuses on one of the alternative modes of assessment which is “the IPA model”. The latter is one of the alternative assessment methods that can play an important role in improving language learning and learners' language skills proficiency. It also explains in details the different phases of this assessment model and how it can be implemented in EFL classrooms in higher education.

Correspondingly, and in light of this study's findings, this study offers a set of pedagogical implications dedicated to researchers interested in language assessment. These implications are summarized as follows:

- The IPA can be an effective method through which assessment can be both an evaluation and a teaching tool. However, EFL teachers should develop a good understanding of the IPA before implementing it. To do so, the IPA manual could be the best support in order to properly introduce the IPA to the students before implementing it in their classrooms. Teachers should have a full understanding of all the phases and how this model can help to assess and enhance their performance, skills, and communicative skills. They should be aware of how they would be evaluated and graded as well.

- Teachers should prepare a full plan of all the phases and subsequent tasks that will be conducted in the EFL classrooms by starting up with determining the objectives of each phase as well as the overall objectives of all the integrated IPA method. Thus, they should be aware that this assessment model is a complex process that implies good preparation, some challenges, and consumes time.

- Teachers should consider that the IPA is both a teaching and an assessment tool. This implies that teaching and assessment are integrated in a seamless process.

- Teachers should consider that the IPA includes both the formative and the summative types of assessment: the formative assessment, involves all tasks and performance activities throughout the unit of instruction, for example: quizzes, writing tasks, reading comprehension tasks, listening comprehension, and speaking tasks. This type provides an opportunity for the students to recognize their weaknesses and flaws with the help of the constant feedback then try to enhance them in summative assessment. Then, the summative assessment includes the presentational tasks of IPA at the end of the unit.

- During the feedback sessions, students can be assisted to recognize, learn, and adopt strategies to develop their skills such as: the reading strategies, including skimming for idea and scanning for details in the first interpretive phase.

- For the interpretive phase, the choice of the authentic text should be appropriate for: the classroom learning activities, the learners' ages, and their linguistic levels taking into consideration that they should modify the tasks to what suit the learners but not the text.

- Before starting each phase, the students should be prepared in a warm up phase through a discussion of the topic or reviewing the previous tasks. This will stimulate their background knowledge and fuel their ideas. Warm-up activities can lower the affective filter and provide students with enough time to be engaged in the course.

- Modelling is a helpful tool, so, at the beginning of each phase teachers can provide the learners with models of the expected performance to prepare them for the tasks.

- In the interpersonal phase, all students should participate in the open discussion and each one of them should speak and interact with the others. The proficiency level disparity, while doing the discussions, will lead low-level students to work harder and correct their flaws with the help to the other classmates.

- Provide feedback to each student right after the presentation. In this way, they will be provided with opportunities to evaluate their own impact and work. The feedback should be co-constructed by the teacher, the one who presented, and the classmates.

- One means of providing and facilitating feedback can be through the use of videotapes. Recording the students' presentations and then make them watch the videos can provide a feedback loop to help them evaluate their own work.

- Addressing students' personal interests while selecting the themes and forming the tasks would fuel their motivation and engagement in the IPA. Thus, teachers can engage students before the start of the course in a discussion in which they express their preferences regarding the topic and themes to be tackled during the whole course. Providing the students with an opportunity to choose different topics in presentational phase related to the theme, for example, could be a good way to vary the presentations and avoid boredom.

- Teachers can provide guidance and support to help the students overcome their speaking weaknesses and challenges including stress, shyness, and being afraid to participate throughout the interpersonal and presentation phases to make confident learners.

- The use of the IPA should be in line with the educational objectives and the curriculum needs. Thus, teachers can adjust the contents and the tasks of the IPA plan they are using.

- This model may not be appropriate to be used in other modules except the oral and written expression modules since they are the only modules that integrate all skills and involve the students in discussions and debates on free topics. Assessment in these two subjects is dedicated to test levels of proficiency in the productive skills through evaluating written or oral products by the learners in forms of oral presentations or essays.

- Seminars and study days on language teaching\ learning and assessment practices can be good occasions to raise awareness about either the IPA or other innovative assessment tools that can serve as alternatives to not just assess but also to develop learners' skills.

5.6 Limitations and Challenges of the IPA

The current study, along with previous research on the implementation of the Integrated Performance Assessment (IPA) in language classrooms, has identified several limitations. These limitations may hinder the success of this assessment model and pose difficulties for instructors. Consequently, it is

imperative to consider the following limitations drawn from this study and prior research when utilising the IPA.

The Learners will not be able to accomplish the subsequent tasks unless they completely understand the listening and reading tasks of the interpretive phase and meet the standards. Therefore, scores and results of the interpersonal and the presentational phases rely on the first phase (Madison, 2019; Troyan, 2007).

Additionally, if the learners lack the proficiency and the skills needed to accomplish one of the phases, the teacher has to spend more time in enhancing the learners' level until they meet the standards and criteria of the given task (Troyan, 2007). This may be an extra pedagogical burden for the teacher that can be a daunting and a time consuming task.

Moreover, language teachers may encounter another challenge in implementing the IPA, which is related to the integration of technology for recording, presenting, and assessing language performance. This challenge arises from the fact that not all classrooms could have access to technology and logistics including laboratories and data shows.

Furthermore, some learners may not be compliant to participate in non-rehearsed task in the interpersonal phase mainly in the oral/speaking tasks. Thus, teacher may have difficulty to encourage them to be involved in conversations and interviews (Tsou, 2005).

Besides, the individual differences among the learners, within the same classrooms, may grow into a serious problem in the implementation of IPA. Thus,

different learners with different levels will not reach the standards equivalently (Troyan, 2007).

Another limitation that can have a serious impact of the implementation of the IPA is teachers' lack of assessment literacy and skills. According to Zapata (2016), the actual failure of the IPA lies in the fact that teachers are not ready for change since they lack a full understanding of this innovative assessment principles, purposes, and actual implementation in their classrooms (Zapata, 2016).

Other challenges that may be encountered by the teachers throughout the implementation of the IPA are similar to those of the communicative approach including, overcrowded classrooms where the involvement of all the learners becomes difficult, the limited resources, as well as time constraints that may be contrasted to the time-demanding nature of the IPA (Madison, 2019).

Research on the area of washback of the IPA should consider using longitudinal designs. Longer engagements in the field could reveal possible impact IPA on teachers' practices and cognitions over the long-term and/or change over time, and even across instructional contexts.

Methodological limitations of this study can be addressed in the intended similar research. These limitations include the limited time available during the second semester. Having more time could have been beneficial in obtaining more dependable and reliable findings.

Another issue faced by the researcher is the lack of resources and similar studies in the field. The implementation of the IPA may vary among studies, and

some studies may have focused on only one skill, making it challenging to compare the results across different studies.

Additionally, a gap in the assessment timeline may have affected the results since the students completed the oral presentational task at the mid-semester and the written presentational task at the end of the semester. Therefore, the study may have been limited by time constraints, preventing a more in-depth exploration of the impact of the IPA on language proficiency.

The generalisability of the research findings could be also one of the limitations. Since the present study only focused on first-year students in one specific course at one university. This may limit the generalisability of the findings to other contexts. Also, it is difficult to assume that the obtained results, from the implementation the IPA model in the oral expression course during the first semester-long, were all accurate and reliable, since the EFL students, who participated in the experiment, studied several other courses, which could also help in improving their proficiency levels. However, this new assessment improved its efficacy in rising students' motivation and developing their positive attitudes towards the oral expression session and towards assessment. Therefore, the IPA model can serve as a language teaching\learning support and an assessment tool as well.

It is important to consider the limitations of the present study as well as the previous studies to ensure that future research addresses these limitations to provide a more comprehensive understanding of the effectiveness of the IPA in enhancing language proficiency.

5.7 Suggestions for Future Research

Recognising the limitations of the present investigation that would, in one way or another, limit the generalisability of the findings, it suggested that subsequent replicated studies would preferably consider the following recommendations:

- Exploring the effectiveness of using the IPA in a longitudinal study is required in order to track language learners' language proficiency levels in a more in-depth exploration of the impact of the IPA on language proficiency. It may also help to gain a profound understanding of the deeper components and processes of the IPA.
- More precise specifications of the study scope variables would be interesting. This involves specifying and limiting the scopes to explore the use of the IPA to improve the communicative skills or only one of the skills such as: the use of this assessment model in the oral expression course to enhance the speaking skill, or in the written expression session to improve the writing skill since this study revealed that these two skills are significantly affected by the IPA.
- Necessary expansion of the study scope to include other academic and educational levels, larger samples, and other populations as well as in other academic contexts.
- Further research can be devoted to study and analyse the co-constructed approach to feedback provided during the IPA and how this helps learners to improve language performance and skills. Further research is recommended on the washback effect of this assessment method on both teaching and learning.
- Experimental studies, testing the effectiveness of using the IPA as a pedagogical and assessment tool in improving the learners' communicative competence, are recommended.

- Studies on the washback of the IPA on both teaching and learning processes are recommended.
- More explorations are needed on testing the validity and the reliability of IPA as a proficiency test.

Conclusion

This chapter provided comprehensive discussions for the findings of the present study within a theoretical framework. These findings, align with the related literature, have confirmed the effectiveness of the use of the IPA model in assessing and enhancing EFL learners' language proficiency, as well as boosting their motivation in the oral expression session.

Further, the discussion of the results in this chapter identified the study outcomes' and pedagogical implications, highlighting the potential benefits of implementing the IPA model in language teaching and assessment. The researcher spotted the limitations encountered during the study, such as the limited sample size and the focus on a single language skill.

To address these limitations and the gaps of the present study as well as in the related literature, this chapter has provided a number of recommendations for further research. These recommendations focus on expanding the study to include a larger and more diverse sample, exploring the use of the IPA model in different language skills and contexts, and investigating the effectiveness of the model over a longer period of time.

Overall, all the findings presented in this chapter have significant implications for foreign language educators, researchers, and policymakers, by providing them with insights into the potential benefits of the implementing the IPA model in language teaching and assessment, as well as identifying areas for future research and development.

General Conclusion

Assessment is a crucial element in education. There is no reference to teaching and learning Foreign languages with no reference to assessment practices. It is a critical tool that helps educators to evaluate the effectiveness of their teaching methods and to identify areas of improvement. The integrated performance assessment (IPA), as an innovative assessment model, is a powerful tool that aims not just to assess and evaluate learners' skills but also to enhance their language proficiency across the four skills of listening, speaking, reading, and writing. The IPA provides learners with opportunities to use the language in real-world situations, promoting a communicative approach to language instruction that emphasizes the development of fluency and accuracy in language use. By integrating language skills in meaningful and authentic ways, the IPA helps learners to develop a more comprehensive and relevant understanding of the language, enabling them to communicate more effectively in a range of settings.

This study represents one of the initial attempts in the Algerian context to address the implementation of the IPA in higher education to help enhancing students' language skills proficiency. Initiated by new questions and taking a different perspective to examine the its effectiveness to improve not only one skill, as in previous studies, but all the four skills together. This study also digs deeper into exploring the washback effect of this assessment tool on learners' attitudes motivation. It sheds light on the importance of implementing such innovative assessment methods as alternatives to the traditional ones. This study carries an enquiry into the process, factors and outcomes of IPA implementation. In this regard, it provides an insightful guide to the understanding of this newly developed assessment model along with its three phases.

In this respect, this study is fueled by the goals of exploring the IPA model in the Algerian higher education context, raising attention to a necessary shift from traditional assessment methods in the Oral Expression module in the Algerian English department to alternatives that are more authentic and that helps learners to improve their language skills and abilities. It urges to draw a holistic and cohesive picture to the IPA model as an alternative in this course and its implementation in EFL classrooms. In doing so, a mixed methods approach was opted for this study whereby an experimental scope was intervened.

Within this scope, a number of research questions were asked targeting the current assessment practices among the Algerian EFL teachers to assess their learners' language learning in the Oral Expression module, the effectiveness of using of the IPA model to enhance EFL students' language skills proficiency, and the washback effect of IPA use on the experimental group.

Given the nature of the adopted research approach and the study, a triangulation of research methods and instruments was used. A number of methodological procedures and different data collection tools were taken including the adoption of an experimental deductive perspective in collecting the targeted data as well as the use of survey research in the form of questionnaires. In the first stage of the study, the researcher used, an online questionnaire designed and administered via google forms dedicated to Algerian EFL university teachers. In the second stage of the study, the main research instrument was a pre-test and post-test experimentation. The tool of the final stage involved another online questionnaire for the experimental group participants of this study.

Each of the three phases constituted a full study that has its own research questions, objectives, and data collection tool. But firstly, a piloting grounding phase was conducted to test the efficacy of the selected tools and the validity of the IPA as an assessment tool.

Then, the first stage was conducted in order to understand the setting of the study, which is the Oral expression course and its assessment practices whereby the IPA is to be implemented as the treatment of the experiment. This stage served as the main platform to conduct this study.

The second phase, as the main study of this research, constituted the treatment that is: the implementation of the IPA. The experiment started with a pre-test in the form of a proficiency test that was adapted from the IELTS academic test (USA). The researcher employed it to test students' English language proficiency variation after a semester-long of using the IPA. Then, the treatment lasted a semester-long, in the oral expression module, through which six sessions were devoted to the implementation of the IPA. Each session lasted for two hours. During these sessions, students were integrated into three modes of communication (interpretive, interpersonal, and presentational) followed by a feedback session after each phase. The teacher used various tasks to assess and boost EFL students' language proficiency. In addition to these six sessions, two supplementary sessions were added. The first one was added at the beginning of the interpersonal phase to refresh students' minds after the spring vacation, and the second was to compensate for the limited and inadequate time because of the large number of students in each group. At the end of this stage and to test the

effect of the IPA, another IELTS proficiency test was conducted as a post-test to compare the results.

The third stage, which is the final step in the present study, as it has been mentioned earlier, involved the use of a questionnaire, devoted to the students of the experimental group who participated in the experiment, to test the washback effect of the IPA on their attitudes and motivation.

The findings obtained from the first stage provided a great deal of information about the current assessment practices in the oral expression module, the challenges that teachers face to assess their students' skills in this course, and their needs. Additionally, this phase revealed that Algerian EFL teachers are ready for a change in this module from the traditional assessment practices toward a more authentic one. Thus, the researcher suggested the IPA as an alternative to traditional methods.

This study revealed significant information about the effectiveness of using the IPA to enhance learners' language skills. Among the important findings include that the implementation of the IPA assessment model can contribute to the promotion of language skills proficiency levels among EFL students. The findings support the use of the IPA assessment model as a valid and effective assessment method as well. Therefore, this study highlights the importance of incorporating innovative and effective assessment strategies into language teaching and learning practices such as the IPA.

Findings attained from the ultimate third stage, reflected the highly positive attitudes of the students towards the implementation of the IPA in their oral

expression course. they also hold highly positive attitudes towards the teacher's feedback and peers' feedback provided in the IPA. Additionally, finding in this phase depicted students' acceptance and positive attitudes toward the integration of reading and writing skills into the oral expression session and they assumed the fact these integrated skills helped them to overcome their weaknesses. Not just that, but the finding also from this questionnaire revealed the positive impact of the IPA on students' motivation.

This study ends up evoking more research in this area and suggesting more gaps to fill in, including the implementation of this assessment model in other courses as in the written expression course, exploring the effectiveness of using the IPA to enhance language learners' communicative skills, exploring its effectiveness in improving learners' speaking skill in the oral expression course. Investigating its implementation and effectiveness in developing learners' language skills proficiency in other contexts and other different levels or even in private schools of foreign languages. It also raises awareness of the need for alternative assessment methods in the oral expression course in higher education, as well as teachers' readiness for this change.

References

- Adair-Hauck, B., & Troyan, F. J. (2013). A descriptive and co-constructive approach to integrated performance assessment feedback. *Foreign Language Annals*, 46, 23–44.
- Adair-Hauck, B., Glisan, E. & Troyan, F. (Eds.) (2014). *Implementing integrative performance assessments* (2nd ed.). Alexandria, VI: The American Council on the Teaching of Foreign Languages. Applied linguistics (3rd Ed.). London: Longman.
- Adair-Hauck, B., Glisan, E. W., & Troyan, F. J. (2013). *Implementing integrated performance assessment*. The American Council on the Teaching of Foreign Languages, Alexandria, VA. 978-0-9896532-4-4.
- Adair-Hauck, B., Glisan, E. W., Koda, K., Sandrock, S. P., & Swender, E. (2006). The integrated performance assessment (IPA): Connecting assessment to instruction and learning. *Foreign Language Annals*, 39, 359–382.
- Ahmad, ShariqueL Wasim, Saeeda & Irfan, Sumaiya & Gogoi, Sudarshana & Srivastava, Anshika & Farheen, Zarina. (2019). Qualitative v/s Quantitative Research. 6. 2828-2832. 10.18410/jebmh/2019/587.
- Alderson, J. C. (2000). *Assessing Reading*. Cambridge: Cambridge University
- Altstaedter, L. L., and Krosl, M. (2018). "Perceptions of Integrated Performance Assessments Among Beginning Spanish College Students: A Preliminary study." *Journal of Foreign Language Teaching and Applied Linguistics* 5.2: 99-114.
- American Council on the Teaching of Foreign Languages (ACTFL). (2012). *ACTFL proficiency guidelines—Speaking, writing, listening and reading* (3rd ed.). Alexandria, VA: Author. Google Scholar

Andrade, H.G. (2000). Using rubrics to promote thinking and learning. *Educational Leadership*, 57(5): 13- 18.

Annenberg/Corporation for Public Broadcasting (CPB). (2004). *Teaching foreign languages K-12: Assessment strategies*. Retrieved from <http://www.learner.org/channel/libraries/tfl/assessment/index.html>.

applied linguistics (3rd Ed.). London: Longman.

Bachman, L. F. and A. S. Palmer. 1996. *Language Testing in Practice: Designing and Developing Useful Language Tests*. Oxford: Oxford University Press.

BENNETT, R. E. (1999) *Reinventing Assessment: Speculations on the Future of Large-Scale Educational Testing* (Princeton, NJ: Educational Testing Service).

BENNETT, R. E. (1999) *Reinventing Assessment: Speculations on the Future of Large-Scale Educational Testing* (Princeton, NJ: Educational Testing Service).

Black, P. and D. Wiliam (1998). "Assessment and classroom learning." *Assessment in Education* 5(1): 7 - 68.

Black, P. and D. Wiliam (1998). "Assessment and classroom learning." *Assessment in Education* 5(1): 7 - 68.

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability* (formerly: *Journal of Personnel Evaluation in Education*), 21(1), 5-31.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the Black Box: Assessment for Learning in the classroom. *Phi Delta Kappan*, 86(1), 8-21.

-
- Bose, J., & Rengel, Z. (2009). *A model formative assessment strategy to promote student-centered self-regulated learning in higher education. US-China Education Review, 6(12), 2935.*
- Boud, D., & Molloy, E. (2013). Rethinking models of feedback for learning: The challenge of design. *Assessment and Evaluation in Higher Education, 38(6), 698-712.*
<http://dx.doi.org/10.1080/02602938.2012.691462>
- Brink, C. S., "A Historical Perspective of Testing and Assessment Including the Impact of Summative and Formative Assessment on Student Achievement" (2011). University of Southern Mississippi. Dissertations. 409
- Brown, D. H. (2000). *Teaching by principles: an interactive approach to language pedagogy* (2nd ed.). London: Longman.
- Brown, H. D. (2001). *Teaching by principles: An interactive approach to language pedagogy.* (2 nd ed.). London: Longman.
- Brown, H. D. (2004). *Language assessment: Principles and classroom practices.* White Plains, NY: Pearson Education.
- Brown, H., (2007). *Teaching by principles.* New York: Pearson, Longman.
- Canale, M. 1983. 'On some dimensions of language proficiency' in J. Oller (ed.). *Issues in Language Testing Research.* Rowley, MA: Newbury House.
- Carless, D., Salter, D., Yang, M., & Joy, L. (2011). *Developing sustainable feedback practices. Studies in Higher Education, 36(4), 395-407.*
<https://doi.org/10.1080/03075071003642449>

-
- Cauley, K. M., & McMillan, J. H. (2010). Formative assessment techniques to support student motivation and achievement. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, 83(1), 1-6. doi: [10.1080/00098650903267784](https://doi.org/10.1080/00098650903267784)
- Chen, Y. S. (2007). Students' changing views and the integrated-skills approach in Taiwan's EFL college classes. *Asia Pacific Education Review*, 8, 27-40. [10.1007/BF03025831](https://doi.org/10.1007/BF03025831).
- Cheriet, H. & Houadjli, A. (2022). EFL Teachers' Attitudes towards the Use of Performance-based Assessment in the Algerian Middle Schools. *El Jamie*, 7, 1586-1610.
- Ciofalo, J., & Wylie, C. E. (2006). Using diagnostic classroom assessment: one question at a time. *Teachers College Record*, 108(1).
- Cizek, G. J. (2010). *An introduction to formative assessment: History, characteristics, and challenges*. In H. L. Andrade & G.J. Cizek (Eds.), *Handbook of formative assessment* (pp. 3-17). New York: Routledge.
- Cloud, N., Genesee, F., & Hamayan, E. (2000). *Dual language instruction: A handbook for enriched education*. Boston: Heinle & Heinle.
- Cognitive/academic language proficiency, linguistic interdependence, the optimum age question and some other matters'. *Working Papers on Bilingualism* 19: 121-9.
- Council of Europe, (2020), *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume*, Council of Europe Publishing, Strasbourg, available at www.coe.int/lang-cefr.

- Council of Europe, (2001), A Common European Framework of Reference for Language Learning and Teaching. Cambridge: Cambridge University Press. Cummins, J. 1979.
- Council of Europe, 2022, <https://www.coe.int/en/web/common-european-framework-reference-languages/table-1-cefr-3.3-common-reference-levels-global-scale>
- Cowie, B. and B. Bell (1999). "A model for formative assessment in science education." *Assessment in Education* 6(1): 101-116.
- Craven, E. (2012). *The quest for IELTS 7.0: Investigating English language proficiency of international students in Australian universities. IELTS Research Reports, 13, 1–61.*
- Creswell, J. W. (1994). *Research design: Qualitative & quantitative approaches*. Thousand Oaks, Calif: Sage Publications.
- Cummins, J. 2008. 'BICS and CALP: empirical and theoretical status of the distinction' in B. Street and N. Hornberger (eds.). *Encyclopedia of Language and Education*. New York, NY: Springer Science + Business Media LLC.
- Danasaputra, I. R. (2003). *Dialogue vs Role Play Techniques. Paper presented at 51th TEFLIN International Conference, Bandung, Indonesia.*
- Davies, A. (1982). "Language Testing Parts 1 and 2." In V. Kinsella (Ed.), *Cambridge Surveys 1* (pp. 127–59). Cambridge, England: Cambridge University Press.
- Davies, A. (2013). *Fifty Years of Language Assessment*. Retrieved from <https://onlinelibrary.wiley.com/doi/full/10.1002/9781118411360.wbcla127>
- Davin, K. J. (2013). Integration of dynamic assessment and instructional conversations to promote development and improve assessment in the language classroom. *Language Teaching Research* 17 (3), 303-322.

Davin, K., Troyan, F. J., Donato, R., & Hellman, A. (2011). Research on the Integrated Performance Assessment in an early foreign language learning program. *Foreign Language Annals*, 44(4), 605-625.

Dawadi, S., Shrestha, S., & Giri, R. A. (2021). Mixed-Methods Research: A Discussion on its Types, Challenges, and Criticisms. *Journal of Practical Studies in Education*, 2(2), 25-36

Diani, A. (2005). Some techniques to make shy students active. Paper presented at LIA International conference, Jakarta.

Diaz Maggioli, Gabriel. (2020). Integrated Performance Assessment. 10.1007/978-3-030-42269-1_3.

Dooley, P. (1999). An investigation into the predictive validity of the IELTS Test as an indicator of future academic success. In K. Martin, N. Stanley & N. Davison (Eds.), Teaching in the disciplines/ learning in context (pp. 114–118). Proceedings of the 8th Annual Teaching Learning Forum, The University of Western Australia, February 1999. Perth: UWA. Retrieved from <http://lsn.curtin.edu.au/tlf/tlf1999/dooley.html>

DUPONT, A., HELD, J. A., HOSSAIN G. S., (2020). Let's Talk about It: Strategies for Integrating Writing and Speaking in the Classroom. ENGLISH TEACHING FORUM. ISSN-1559-663X.(22_31).

Entwistle, N.J. & Tait, H. (1990). Approaches to learning, evaluations of teaching, and preferences for contrasting academic environments. *Higher Education*, 19, 169-194

- Espinosa, L. F. (2015). Effective Use of Performance-based Assessments to Identify English Knowledge and Skills of EFL Students in Ecuador. *Theory and Practice in Language Studies*, 5(12), 2441–2447. <https://doi.org/10.17507/tpls.0512.02>
- Feast, V. (2002). *The impact of IELTS scores on performance at university. International Education Journal*, 3(4), 70–85.
- Gay, L.R., Geoffrey, E. M., Airasian P. W.(2012). Educational Research: Competencies for Analysis and Applications. USA: Pearson Education.
- Gharbavi, Abdullah & Seyyed, & Mousavi, Seyyed Ahmad & Bailey, R & Seetharaman, S & Gavin, C & Shukla, N. (2012). Do Language Proficiency Levels Correspond to Language Learning Strategy Adoption? . 5. 10.5539/elt.v5n7p110.
- Gibbs, G. (2006). How assessment frames student learning. In C. Bryan & K. Clegg (Eds.), *Innovative assessment in higher education* (pp.23-36). London: Routledge.
- Glisan, E. W., Adair-Hauck, B., Koda, K., Sandrock, S. P., & Swender, E. (2013). *ACTFL integrated performance assessment*. Yonkers, NY: ACTFL.
- Glisan, E. W., Adair-Hauck, B., Koda, K., Sandrock, S. P., & Swender, E. (2003). *ACTFL integrated performance assessment*. Yonkers, NY: ACTFL.
- Glisan, E. W., Uribe, D., & Adair-Hauck, B. (2007). Research on integrated performance assessment at the post-secondary level: Student performance across the modes of communication. *Canadian Modern Language Review*, 64, 39-68.
- Grabe, W. (1991). *Current developments in second language reading research. TESOL Quarterly*, 25(3), 375-406.<http://dx.doi.org/10.2307/3586977>

- Green, A. (2005). *EAP study recommendations and score gains on the IELTS Academic writing test. Assessing Writing, 10, 44–60. doi:10.1016/j.asw.2005.02.002*
- Greenstein, L. (2010). What teachers really need to know about formative assessment. USA: ASCD.
- Gurel, D. K., Eryilmaz, A., & McDermott, L. C. (2015). A review and comparison of diagnostic instruments to identify students' misconceptions in science.
- Hamrayevna, R., Rashidovna, S., & qizi, R. (2019). Impact of extensive reading to develop speaking skills. *International Journal on Integrated Education* 2(4):11-13. DOI:[10.31149/ijie.v2i4.98](https://doi.org/10.31149/ijie.v2i4.98)
- Hoerr, Thomas R. 2000. *Becoming a Multiple Intelligences School*. ASCD
- Humphreys, Pamela & Haugh, Michael & Fenton-Smith, Ben & Lobo, Ana & Michael, Rowan & Walkinshaw, Ian. (2012). *Tracking international students' English proficiency over the first semester of undergraduate study. IELTS research reports. 1. 1-41.*
- Hungyo, J and Kijai, J. (2009). *The Effect of Integrated and Segregated Skills Approach on English Language Acquisition among Freshmen Nursing and Business Students in Mission College Thailand. The Scriptor, Journal of Arts & Humanities Department. Thailand.*
- J. L. Herman. P. R. Aschbacher, L. Winters. (1992). A practical guide to alternative assessment. Alexandria, VA: Association for Supervision and Curriculum Development.
- James, M. (2010) An overview of Educational Assessment. In: P. Peterson, E. Baker & B. McGaw (Eds.) *International Encyclopedia of Education. Vol.3: 161-171. Oxford: Elsevier.*

-
- Kern, R. (2000). *Literacy and language teaching*. New York: Oxford University Press.
- Kissau, S., & Adams, M. J. (2016). Instructional design making and IPAs: Assessing the modes of communication. *Foreign Language Annals*, 49(1), 105-123.
- Klimova, Blanka. (2015). The Role of Feedback in EFL Classes. *Procedia - Social and Behavioral Sciences*. 199. 172-177. 10.1016/j.sbspro.2015.07.502.
- Knight, P. (1995). *Assessment for learning in higher education*. London: Kogan Page.
- Krashen, S. (1985). *The Input Hypothesis Issues and Implications*. London: Longman.
- Krashen, S. D., & Terrell, T. D. (1989). *The Natural Approach: Language Acquisition in the Classroom*. Pergamon: PrenticeHall.
- Kurniawan, I., Sabaruddin, S., Lubis, A.A. (2022). The implementation of integrated performance assessment (IPA) model to improve English students' speaking skills. *JOALL (Journal of Applied Linguistics and Literature)*, 7(2), 432-443. <https://doi.org/10.33369/joall.v7i2.20860>
- Lammers, W. J., & Badia, P. (2004). *Fundamentals of behavioral research*. Belmont, Australia: Wadsworth Publishing.
- Larsen-Freeman, D., & Anderson, M. (2013). *Techniques and principles in language teaching (3rd ed.)*. Oxford university press.
- Learning & Teaching, Centre Learning Skills Group. (2013). *Developing Your Students' English Language Proficiency The Academic Literacies Approach*. Macquarie University.
- Lee, K., & Lim, J. (2020). The effectiveness of using the Integrated Performance Assessment (IPA) in improving English proficiency of Korean learners. *Journal of*

Educational Evaluation for Health Professions, 17, 1-9.
<https://doi.org/10.3352/jeehp.2020.17.7>

Lisa A. Harris et al. (2015) *Implementing the Foreign Language Standards of Learning in Virginia Classrooms: A Guide for Teachers*. The Virginia Department of Education

Madison, Stephanie. (2019). Integrated Performance Assessments: A review of the literature and steps to move forward. 5. 99-111.

Marguerite M. Clarke , George F. Madaus , Catherine L. Horn & Miguel A. Ramos (2000): Retrospective on educational testing and assessment in the 20th century, *Journal of Curriculum Studies*, 32:2, 159-181

Martel, J. (2018). Postsecondary students' and instructors' evaluative comments about ACTFL's Integrated Performance Assessment. *Journal of Applied Language Learning*, 28(1), 1–18.

Martel, J., & Bailey, K. M. (2016). Exploring the trajectory of an educational innovation: Teachers' attitudes toward summative IPA implementation in an intensive summer language program at the post-secondary level. *Foreign Language Annals*, 49(3), 530-543.

Maureen Snow Andrade (2009) The Effects of English Language Proficiency on Adjustment to University Life, *International Multilingual Research Journal*, 3:1, 16-34, DOI: [10.1080/19313150802668249](https://doi.org/10.1080/19313150802668249)

McKay, P. (2006). *Assessing young learners*. Cambridge: Cambridge University Press

McKay, P. (2006). *Assessing young learners*. Cambridge: Cambridge University Press.

- Masuku, Mfundo & Jili, Nokukhanya & Sabela, Primrose. (2020). Assessment as A Pedagogy and Measuring Tool in Promoting Deep Learning In Institutions of Higher Learning. *International Journal of Higher Education*. 10. 274. 10.5430/ijhe.v10n2p274.
- Mertler, C. A. (2001). Designing scoring rubrics for your classroom. *Practical Assessment, Research and Evaluation*, 7(25).
- Murdibjono. (2001). *Conducting small group discussion. Bahasa dan Seni*.19(1), 139-150.
- National Standards in Foreign Language Education Project (NSFLEP). (2006). *Standards for foreign language learning in the 21st century*. Lawrence, KS: Allen Press.
- Nicol, D., & Macfarlane-Dick, D. (2006). *Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. Studies in Higher Education*. 31(2), 199-218. <https://doi.org/10.1080/03075070600572090>
- O'Loughlin, K. & Arkoudis, S. (2009). *Investigating IELTS exit score gains in higher education. IELTS Research Reports*, 10(3), 1–90.
- Omaggio, A. C. (1986). *Helping learners succeed: Activities for the foreign language classroom*. Washington: Center for Applied linguistics.
- Panadero, E., and A. Jonsson. (2013). "The Use of Scoring Rubrics for Formative Assessment Purposes Revisited: A Review." *Educational Research Review* 9: 129–144. doi:10.1016/j.edurev.2013.01.002.
- Peregoy, S. F. & Boyle, O. F. (2001). *Reading, Writing, & Learning in ESL: A Resource Book for K-12 Teachers*. New York: Addison Wesley Longman, Inc

- Popham, W. J. (2008). *Classroom assessment: What teachers need to know (5th ed.)*. Boston: Prentice Hall.
- Povey E. (2019). *IPA: A Framework for Language Teaching Based on Bloom's Taxonomy*. *IJO - INTERNATIONAL JOURNAL OF EDUCATIONAL RESEARCH*. 02 .12. https://www.edwardtesol.com/uploads/2/6/5/8/26581579/264-article_soft_copy.pdf
- Prabhu, N. S. (1987). *Second language pedagogy*. Oxford: Oxford University Press.
- Prosser, M. & Webb, C. (1994) *Relating the Process of Undergraduate Essay Writing to the Finished Product*. *Studies in Higher Education*, 19:2 pp.125-138. <https://doi.org/10.1080/03075079412331381987>
- Rea, Jean B., (2010) *You Say Ee-ther and I Say Eyether: Clarifying Assessment and Evaluation*, Retrieved from <https://www.td.org/newsletters/atd-links/you-say-ee-ther-and-i-say-eye-ther-clarifying-assessment-and-evaluation#gsc.tab=0>
- Rea, Jean B., (2010) *You Say Ee-ther and I Say Eyether: Clarifying Assessment and Evaluation*, Retrieved from <https://www.td.org/newsletters/atd-links/you-say-ee-ther-and-i-say-eye-ther-clarifying-assessment-and-evaluation#gsc.tab=0>
- Richards, J. C., Platt, J., & Platt, H. (1992). *Longman Dictionary of Language Teaching and Applied Linguistics*. London: Longman.
- Richards, Jack C and Willy ARenandya. (1990). *Methodology in Language Teaching*. Cambridge: Cambridge U.
- River, W. M. & Temperly, M. S. (1999). *A Research on Language Skills*. Cambridge: Cambridge University Press.

- Rogers, J., & Révész, A. (2019). Experimental and quasi-experimental designs. *The Routledge Handbook of Research Methods in Applied Linguistics*.
- Rudner, L. M., & Boston, C. (1994). Performance assessment. *ERIC Review*, 3(1), 2–12.
- Ruiz-Primo, M. A., & Furtak, E. M. 2007. Exploring Teachers' Informal Formative Assessment Practices and Students' Understanding in the Context of Scientific Inquiry. *Journal of Research in Science Teaching*, 44 (1): 57-84.
- Ruiz-Primo, M. A., & Furtak, E. M. 2007. Exploring Teachers' Informal Formative Assessment Practices and Students' Understanding in the Context of Scientific Inquiry. *Journal of Research in Science Teaching*, 44 (1): 57-84.
- Sadler, D. (1998). "Formative assessment: revisiting the territory." *Assessment in Education* 5(1): 77-84.
- Syaifuddin, M. (2019). The effect of students' perception on classroom assessment to students' attitudes. *Journal of Physics: Conference Series*, 1280(4).
<https://doi.org/10.1088/1742-6596/1280/4/04202>
- Sambell, K., McDowell, L., & Montgomery, C. (2013). *Assessment for Learning in Higher Education*. Abington: Routledge.
- SARNOU H., KOÇ S., HOUCINE S. & BOUHADIBA F. (2012), LMD New System in the Algerian University. *Arab World English Journal*, pp. 179 – 194.
https://www.academia.edu/29055367/LMD_New_System_in_the_Algerian_University
- Sedor, N. (2022). Integrated Performance Assessment (IPA): Implementation, task types, and feedback. *Foreign Language Annals*, 1–21.
<https://doi.org/10.1111/flan.12661APP>

- Shankar L., (2016) Differences between Testing, Assessment, and Evaluation, Retrieved from <https://www.linkedin.com/pulse/differences-between-testing-assessment-evaluation-lahiri>)
- Shrum, J. L., & Glisan, E. W. (2010). *Teacher's handbook: Contextualized language instruction. 4th ed.* Boston, MA: Heinle Cengage Learning.
- Skehan, P. (1988). *State of the art article: Language testing part 1. Language Teaching*, 21121.
- Skehan, P. (1989a). *State of the art article: Language testing part 2. Language Teaching*, 113.
- Stanford SRN Informational booklet. (2008). What is the Performance-Based Assessment? Retrieved from <https://edpolicy.stanford.edu/sites/default/files/events/materials/2011-06-linked-learning-performance-based-assessment.pdf>
- Stephenson, A., Johnson, D. F., Jorgensen, A., & Young, M. J., (2003) *Assessing English Language Proficiency: Using Valid Results to Optimize Instruction.* Pearson Inc.
- Stevens, D. D., & Levi, A. J. (2005). *Introduction to rubrics.* Sterling, VA: Stylus.
- Stiggings, R. (2001). *Student-involved classroom assessment.* (3rd ed.). Upper Sadle River, NJ: Merrill-Prentice Hall.
- Suban, T. S. (2021). *Teaching speaking: Activities to promote Speaking Skills in EFL Classrooms. Lectio: Journal of Language and Language Teaching*, 1(1), 41-50
- Surbhi S., (2017), *Difference Between Assessment and Evaluation*, Retrieved from <https://keydifferences.com/difference-between-assessment-and-evaluation.html>
- Teachers of English to Speakers of Other Languages, Inc. (2006). *PreK–12 English language proficiency standards.* Alexandria, VA: Author.

- Tertiary Education Quality and Standards Agency. (2013). 'Quality Assessment: English Language Proficiency. Terms of Reference. <http://www.teqsa.gov.au/for-providers/quality-assessments>
- Troyan, F. (2007). Planning to meet a learning community's needs. *NECTFL Review*, 61, 141-151.
- Troyan, F. (2012). Standards for foreign language learning: Defining the constructs and researching learner outcomes. *Foreign Language Annals*, 45(S1), S118–S140.
- Troyan, F. (2016). Assessing what matters within content-based foreign language teaching through Integrated Performance Assessment. In L. Cammarata (Ed.), *Content-based foreign language teaching: Curriculum and pedagogy for developing advanced thinking and literacy skills* (pp. 147-169). New York, NY: Routledge.
- Troyan, F. J. (2008). Being authentic: Assessing standards-based tasks in a content-based curriculum. *The Language Educator*, 3(4), 52-54.
- Ur, P. (2012). *A course in English language teaching*. Cambridge University Press.
- Widiati, U., & Cahyono, B. Y. (2006). The teaching of EFL speaking in the Indonesian context: The state of the art. *Bahasa dan Seni*, 34(2), 269-292. <http://sastra.um.ac.id/wpcontent/uploads/2009/10/The-Teaching-of-EFL-Speaking-in-the-Indonesian-ContextThe-State-of-the-Art-Utami-Widiati-Bambang-Yudi-Cahyono.pdf>
- Vossoughi, H., & Javaherian, J. (2000). An encyclopedic dictionary of teaching English as a foreign language. Tehran: Rahnama Publication

-
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Vygotsky, L. S. (1986). *Thought and language*. Cambridge: MIT Press.
- Vygotsky, L. S. (1987). *The collected works of Lev S. Vygotsky. Volume I: Problems of general psychology* (R. Rieber & A. Carton, Eds.). New York: Plenum Press.
- Vygotsky, L. S. (1994). The problem of the environment. In R. Van Der Veer & R. Valsiner (Eds.), *The Vygotsky reader* (pp. 338–354). Cambridge: Blackwell.
- Weigle, S. (2002). *Assessing writing*. Cambridge: Cambridge University Press. P46
- White, H., & S. Sabarwal (2014). *Quasi-experimental Design and Methods, Methodological Briefs: Impact Evaluation 8*, UNICEF Office of Research, Florence.
- Wiggins, G. (1989, May). A true test: Toward more authentic and equitable assessment. *Phi Delta Kappan*, 70, 9.
- Wiggins, G. (1989, May). A true test: Toward more authentic and equitable assessment. *Phi Delta Kappan*, 70, 9.
- Wiggins, G., & McThige, J. (2005). *Understanding by design* (expanded 2nd ed.). Alexandria, VA: Association for Supervision and Curriculum Development.
- Wilbrink, B. (1997). *Assessment in historical perspective. Studies in Educational Evaluation*, 23(1), 31–48.
- Wilbrink, B. (1997). *Assessment in historical perspective. Studies in Educational Evaluation*, 23(1), 31–48.

- Wilbrink, B. (1997). *Assessment in historical perspective. Studies in Educational Evaluation, 23(1), 31–48.*
- William, D. (2011). What is assessment for learning? *Studies in Educational Evaluation, 37(1), 3-14.*
- William, D., & Thompson, M. (2007). *Integrating Assessment with Learning: What Will It Take to Make It Work?*
- Wraga, W. (2017). *Understanding the Tyler rationale: Basic Principles of Curriculum and Instruction in historical context. Espacio, Tiempo y Educación. 4. 227. 10.14516/ete.156.*
- Yambi, Tomas & Yambi, Caluyua. (2020). ASSESSMENT AND EVALUATION IN EDUCATION.
- Zapata, G. C. (2016). "University Students' Perceptions of Integrated Performance Assessment and the Connection between Classroom Learning and Assessment." *Foreign Language Annals 49.1: 93-104.*

Appendices

Appendix A

Teachers' Questionnaire

Personal information

1. Please specify your gender

- Male
- Female

2. What is your rank

- Master
- Magister
- Docotrate

3. How long have you been teaching

- 1-5
- 5-10
- More than 10

Part two: teachers' perceptions and practices towards oral expression teaching and assessment

4. In your opinion, what are the main objectives of teaching oral expression?

- To develop students' receptive skills
- To develop students' productive skills
- To develop students' self-confidence.
- To develop their presentations skills
- To enhance their communicative skills.

5. Which skill is most emphasised in teaching oral expression. Rank them from "so important 1 to less important 4"

- Listening
- Speaking
- Reading
- Writing
- All of the above

6. What are the most used activities in your oral expression session?

- Role plays

- Group work
- Movies and songs
- Listening comprehension activities
- Speaking activities
- Other, please specify

7. Which of the above you think students enjoy more?

8. Do you allow your students to suggest activities and topic of their interest?

- a) Yes, all the time
- b) Yes, sometimes
- c) No, I just stick to the syllabus

9. What are the hindrances you encounter when teaching oral expression?

10. Do you integrate the four skills in the oral expression session?

- Never
- Sometimes
- Always

10.1. if you say yes how do you integrate them in the oral expression session?

Assessment of oral expression

11. What types of assessment do you generally use in the oral expression class?

- Role play
- Listening comprehension activities
- Reading comprehension activities
- Written expression activities
- Classroom interaction and discussion
- Oral presentation
- Paper pen tests
- Other, please specify.....

12. How often do you provide Feedback to students during the speaking activities?

- Always
- Sometimes
- Never

13. What type of feedback do you usually provide your students with in the oral expression session?

- Oral feedback
- Written feedback
- Peers feedback
- Others, specify

14. How do you provide feedback to students

- Directly correct the mistake
- Just show them the mistake
- Show the mistake and correct it
- Assist students to find their flaws and correct them
- Others, please specify.....

15. What types of mistakes do you focus on when evaluation oral expression?

- Pronunciation mistake
- Grammatical mistakes
- L1 transfer
- Others, please specify

16. What are the difficulties you face when testing students' skills in oral expression session?

17. Do you integrate the four skills when evaluating oral expression?

- Never
- Yes, sometimes
- Yes, all the time

18. Do you think integrating the four skills in assessing oral expression can be efficient?

19. Do you think that traditional types of assessment are sufficient to test students' language proficiency?

- Yes
- No
- Not sure

20. Do you use alternative assessment in your oral expression class ?

- Yes
- No

20.1 What alternative assessment methods you think serve to assess language proficiency?

Appendix B

IELTS Reading Test

Read the passage carefully.

TOURISTS IN A FRAGILE LAND

Task 1: *Choose the best option A, B, C, or D for each gap.*

1 As a scientist working in Antarctica, I spend most of (1) _____ time in the lab studying ice. I am trying to find out the age of Antarctic ice. All we know for certain is that it is the oldest ice in the world. The more we understand it, the more we will understand the (2) _____ weather of the Earth. Today, as with an increasing number of days, I had to leave my work to greet a group of tourists who were taking a vacation in this continent of ice. And even though I can appreciate their desire to experience this **vast** and beautiful **Landscape**, I feel Antarctica should be closed to tourists.

2 Because Antarctica is the center of important scientific research, it must (3) _____ preserved for this purpose. Meteorologists are now looking at the effects of the ozone hole that was discovered above Antarctica in 1984. They are also trying to understand global warming. If the Earth's temperature continues to increase, the health and safety of every living thing on the planet will be affected. Astronomers have a unique view of space and are able to see it very (4) _____ from Antarctica. Biologists have a chance to learn more about the animals that **inhabit** the coastal areas of this frozen land. Botanists study the plant life to understand how it can live in such a **harsh** environment, and geologists study the Earth to learn more about how it was formed. There are even psychologists who study how people behave when they live and work together in such a **remote** location.

3 When tourist groups come here, they take us away from our research. Our work is difficult, and some of our projects can be damaged by such simple mistakes as opening the wrong door or bumping into a small piece of equipment. Tourists in Antarctica can also hurt the environment. Members of Greenpeace, one of the world's leading environmental organizations, (5) _____ that tourists leave trash on beaches and **disturb** the plants and animals. In a place as frozen as Antarctica, it can take 100 years for a plant (6) _____ back, and tourists can easily damage penguin eggs. Oil spills are another problem caused by tourism. Oil spills not only kill penguins but can also destroy scientific projects.

4 The need to protect Antarctica from tourists becomes even (7) _____ when we consider the fact that there is no government here. Antarctica belongs to no country. Who is making sure that the penguins, plants, and sea are safe? No one is responsible. In fact, we scientists are only **temporary** visitors ourselves. It is true that the number of tourists (8) _____ visit

Antarctica each year is small compared to the number of those who visit other places. However, these other places are inhabited by local residents and controlled by local governments. They have an interest in protecting their natural environments. Who is concerned about the environment of Antarctica? The scientists, to be sure, but not necessarily the tour companies that make money from sending people south.

5 If we don't protect Antarctica from tourism, there (9) _____ be serious **consequences** for us all. We might lose the results of scientific research projects. It's possible that these results could teach us (10) _____ about the causes and effects of climate change. Some **fragile** plants and animals might die and disappear forever. This could damage the balance of animal and plant life in Antarctica. We know from past experience that when things get unbalanced, harmful changes can occur. Clearly, Antarctica should remain a place for careful and controlled scientific research. We cannot allow tourism to bring possible danger to the planet. The only way to protect this fragile and important part of the planet is to stop tourists from traveling to Antarctica.

Task 2: Choose the best option A, B, C, or D for each gap.

11. The word **vast** in paragraph 1 is closest in meaning to

(A) very large (B) pretty (C) small (D) faraway

12. The word **landscape** in paragraph 1 is closest in meaning to (A) an area of land for living

(B) a painting

(C) an icy environment (D) the way an area of land looks

13. The word **preserved** in paragraph 2 is closest in meaning to (A) closed (B) explored (C) protected

(D) changed

14. The word **inhabit** in paragraph 2 is closest in meaning to (A) avoid (B) cover (C) prefer

(D) live in

15. The word **harsh** in paragraph 2 is closest in meaning to. (A) unpleasantly difficult (B) mild (C) dry (D) cold

16. The word **remote** in paragraph 2 is closest in meaning to _____. (A) ideal (B) isolated (C) hostile (D) lonely

17. The word **disturb** in paragraph 3 is closest in meaning to _____. (A) make somebody/something happy

(B) bring somebody/something comfort

(C) annoy somebody/something

(D) cause somebody/something to die

18. The word **temporary** in paragraph 4 is closest in meaning to _____. (A) rich and adventurous (B) careful

(C) lasting for a limited time (D) responsible

19. The word **consequences** in paragraph 5 is closest in meaning to _____. (A) acquaintances (B) effects

(C) causes (D) disasters

20. The word **fragile** in paragraph 5 is closest in meaning to _____. (A) native (B) rare (C) useful (D) easily damaged

(10 points) (20 pts)

Read the passage carefully.

The white face, dark eyes and hair, and blood red lips: both foreigners and the Japanese are fascinated by these beautiful and mysterious women.

Makiko is training to be a geisha. Not many girls want to be a geisha in Japan today. Makiko's parents want her to go to university, study medicine and become a doctor. But Makiko's grandfather paid for her training and bought the kimonos she needed. It's very expensive to become a geisha. You have to have a different kimono for each month of the year, and today a kimono costs three million yen, that's about \$30,000.

It's a hard life for a trainee geisha. **She** has to leave her family and move to a special boarding house called a 'maiko house'. Here, she has to learn traditional Japanese arts such as playing instruments, performing the **tea ceremony**, arranging flowers, singing and dancing. She has to take many difficult tests and exams. Only the best will pass everything and become geishas many years later.

We asked Makiko to describe exactly what a geisha does. 'A geisha has to serve customers and also entertain **them**. She has to sing and dance, and make good conversations.'

Did she enjoy her life as a trainee geisha?

‘I love it. But it’s hard work. Sometimes I get tired of wearing the kimonos and I want to put on a pair of jeans and go to school like a normal teenager. But I can’t have a normal life now. I don’t mind. I feel very lucky.’

And what about later? Can she have a family?

‘Of course. A geisha can have relationships like anybody else and she can get married when she chooses.’ In Japan today there are fewer than a thousand geishas, but they play an important role in preserving Japanese culture and history.

Choose the best option A, B, C, or D.

21. What is the main idea of the passage?

- (A) How to be a famous geisha.
- (B) Interesting life of geishas.
- (C) Life and work of geishas.
- (D) How geishas feel about their life.

22. Why is Makiko training to be a geisha?

- (A) Her parents want her to do so.
- (B) Her grandfather supported her financially.
- (C) She was lucky to be chosen.
- (D) Being a geisha is Makiko’s dream.

23. What is **NOT** a requirement for a geisha?

- (A) She has to pass all difficult tests and exams.
- (B) She can’t stay home during the training.
- (C) She has to learn music.
- (D) She has to play modern instruments.

24. The word **She** in the passage refers to

- (A) Makiko
- (B) A trainee geisha
- (C) The writer
- (D) A family member

25. Why does the writer mention **tea ceremony**?

- (A) To express he/she likes tea.
- (B) To indicate that geishas love tea.
- (C) To give an example of Japanese traditional arts.
- (D) To show that all Japanese drink tea in a special way.

26. The word **them** in the passage refers to

- (A) geishas
- (B) tests and exams
- (C) customers
- (D) conversations

27. Give a title to the passage:

.....

28. Give general Idea to the passage:

PART THREE (10 points) *Read the passage carefully.*

Car theft is a constant problem, but if you own a newer model of car, there is less likelihood of it being stolen. Newer models incorporate security systems that help prevent theft. However, thieves often steal car keys to access newer vehicles so it is a good policy to know where your keys are at all times. Never leave them in your car, even if it is parked in your garage at home. You make life easy for the burglar if you leave keys in a receptacle just inside the door or hanging from a hook. Think of a less obvious place, hidden from view. Once a thief has access to your keys, everything you own is at risk, including your personal safety.

While older cars are a potential target, there are some simple precautions you should take. Definitely, lock your car before you leave it, even for the shortest time, even in the driveway of your house. Think of installing an alarm or a steering lock or consider a mechanical immobilizer that fits across the steering wheel. This is a relatively inexpensive item but it is a good deterrent for thieves. They generally go for the easiest and fastest option.

Obviously, it is in your own interest to make your vehicle as difficult to steal as possible – not only to avoid stress and inconvenience, but also because your insurance premiums will soar if you make a claim. If you are unfortunate enough to have your vehicle stolen, phone the police immediately. You will need to provide information relating to your vehicle including make, model, year, colour, registration number and VIN (vehicle identification number). You need all of this when you file a stolen vehicle report. Police will also log any personal items that were in the vehicle. If these items include identifying documents such as a registration certificate, this will make the illegal resale of the vehicle simpler for the thieves, so keep these documents in a separate place.

Next, contact your insurance company. If your wallet was in the car and it contained credit cards, you will need to notify your bank. If your house keys were in the car, call a locksmith and have your locks changed, and if you are still making payments on the car, you will need to inform the finance company. Both your insurance company and the finance company will ask you for copies of the police report so it is best to obtain one in the first instance.

Complete the notes below.

Choose NO MORE THAN THREE WORDS from the text for each answer.

Write your answers in boxes 8–14 on your answer sheet.

- Modern cars are more difficult to steal because of their **29**.....
- If they have your keys, burglars could steal your car or even threaten your **30**.....
- older models are easier to steal, so get an alarm or steering lock or use an immobilizer (cheap but effective **31**
- The loss of your car is stressful and also makes your **32** more expensive
- If your car is taken:
 - call the police
 - Supply all identifying information for vehicle – fill out a **33**
 - NB: the **34** of the vehicle will be easier for the thieves if you leave official papers inside
 - call the insurance company, bank (if necessary), and the **35** if you owe money on the car

Give synonyms to:

Burglar = 36.....

Vehicle=37.....

.

Illegal=38.....

Locksmith =39.....

40. Give a title to the text.....

Appendix C

IELTS Speaking Test

PART 1 – INTRODUCTION AND INTERVIEW:

Let's talk about your home town or village.

TASK	SCORE	Fluency and coherence	Lexical resources	Gramatical range and accuracy	pronunciation
What kind of place is it?	1	1	1	1	1
What's the most interesting part of your town/village?	1	1	1	1	1
What kind of jobs do the people in your town/village do?	1	1	1	1	1
Would you say it's a good place to live? Why?	1	1	1	1	1

PART 2 – INDIVIDUAL LONG TURN: Candidate Task Card

Describe a well-known person you like or admire.

You should say:

Task	SCORE	Fluency and coherence	Lexical resources	Gramatical range and accuracy	pronunciation
Who is this person?	1	1	1	1	1
What has this person done?	1	1	1	1	1
Why is this person well-known?	1	1	1	1	1
Explain why you like or admire this person	1	1	1	1	1

You will have to talk about the topic for 1 to 2 minutes.

You have one minute to think about what you're going to say.

You can make some notes to help you if you wish.

PART 3 – TWO-WAY DISCUSSION:

We've been talking about a well-known person that you admire and I'd like to discuss with you

a few more questions related to this topic. Let's consider famous people in your country.

TASK	SCORE	Fluency and coherence	Lexical resources	Gramatical range and accuracy	pronunciation
What kind of people become famous in your country?	1	1	1	1	1
• What's different about people who became famous in the past with the people who become famous these days?	1	1	1	1	1
Do you think it was more difficult to become famous in the past?	1	1	1	1	1
• What kind people may become famous in the future?	1	1	1	1	1

IELTS listening score

Correct Answers	Band Scores
39 - 40	9
37 - 38	8.5
35 - 36	8
32 - 34	7.5
30 - 31	7
26 - 29	6.5
23 - 25	6
18 - 22	5.5
16 - 17	5
13 - 15	4.5
11 - 12	4

Appendix D
IELTS Listening

test <https://www.ielts.org/for-test-takers/test-format>

PART 1 Questions 1 and 2 Choose the correct letter, A, B or C. à

1. Type of insurance chosen: A Economy B Standard C
Premium
2. Customer wants goods delivered to: A port B home
C depot
3. Customer wants to pay: A cash B credit card
C check

PART 2 Questions 3 – 5 Answer the questions below.

Write NO MORE THAN THREE WORDS AND/OR A NUMBER for each answer.

4. What TWO factors can make social contact in a foreign country difficult?
.....
5. Which types of community group does the speaker give examples of?
.....
6. In which TWO places can information about community activities be found?
.....

PART 3 Questions 6 – 9 Complete the sentences below.

Write NO MORE THAN TWO WORDS for each answer.

7. Studying with the Open University demanded a great deal of
.....
8. Studying and working at the same time improved Rachel's skills
9. It was helpful that the course was structured in
10. She enjoyed meeting other students at.....

PART 3 Questions 10 - 15

11. What does Jack tell his tutor about each of the following course options?

- A He'll definitely do it. B He may or may not do it. C He won't do it.

Write the correct letter, A, B or C next to questions 21- 25.

You may choose any letter more than once.

12. Media Studies
13. Women and Power
14. Culture and Society

15. Identity and Popular Culture

16. Introduction to Cultural Theory

PART 1 Questions 16– 4

Which hotel matches each description?

Choose your answers from the box and write the correct letter A-E next to questions 1-4.

A The Bridge Hotel

B Carlton House

C The Imperial

D The Majestic

E The Royal Oak

17. is in a rural area

18. only opened recently

19. offers facilities for business functions

20. has an indoor swimming pool

Appendix E
IELTS Writing Test

WRITING TASK:

You should spend about one hour on this task.

Write at least 150 words.

You have seen a job advertised in the newspaper. You want to apply for the job. Write a letter to the manager of the company. In your letter:

- provide information about yourself
 - give details of your previous work experience
 - say why the company should consider you for the job
- You do NOT need to write any addresses.

Begin your letter as follows: Dear ...

Appendix F

The IPA Reading Phase

Text for Reading Comprehension:

21 Top-Rated Tourist Attractions & Things to Do in London

Written by Bryan Dearsley and Shandley McMurray

<https://www.planetware.com/tourist-attractions-/london-eng-1-lon.htm>

Retrieved on : 07/03/2022.

London is one of the world's most fascinating cities... With so many impressive tourist attractions and captivating things to do, it's no wonder London is one of the most visited cities worldwide, receiving over 20 million tourists each year. Britain's bustling capital offers something for everyone, including shopaholics, foodies, adventurers, historians, and kids; but this can make it tough to choose what to do first.

Should you hit one of the top museums (many of which are free to enter), enjoy a picnic in one of the expansive parks, tour a royal palace, or wander around a breathtaking garden? Perhaps you'd prefer to take in a show, ride a horse through a grove, take a flight on the London Eye, or enjoy a traditional afternoon tea at Harrods.

1. Visit Buckingham Palace and Watch the Changing of the Guard

One of Britain's most iconic buildings, Buckingham Palace is also the scene of London's most popular display of pomp and circumstance, the Changing of the Guard. Drawing crowds at 11:30am regardless of the season, this colorful and free display of precision marching and music also takes place at **St. James's Palace**, after which you can follow the band along The Mall as they march between sites.

Buckingham Palace was built in 1837 and has been the London residence of the Royal Family since Queen Victoria's accession. If you're wondering whether the Queen is in, look at the flagpole atop the building: if the royal standard is flying day and night, she's at home. On special state occasions, she and members of the Royal Family may even emerge on the central balcony.

When the Queen's away at her summer palace in Scotland, visitors can purchase tickets for tours of the State Rooms, the Queen's Gallery, and the Royal Mews.



2. See the Royal Jewels at the Tower of London and Walk across Tower Bridge

From prison to palace, treasure vault to private zoo, the magnificent Tower of London has fulfilled many different roles over the centuries. One of Britain's most iconic structures, this spectacular World Heritage Site offers hours of fascination for visitors curious about the country's rich history - after all, so much of it happened here. Inside the massive White Tower, built in 1078 by William the Conqueror, is the 17th-century Line of Kings with its remarkable displays of royal armaments and armor.



The adjacent Tower Bridge, its two huge towers rising 200 feet above the River Thames, is one of London's best-known landmarks (fascinating behind-the-scenes tours are available). Walk across for the best Tower views, as well as a glimpse of London Bridge (which many mistakenly believe Tower Bridge to be) in the distance farther along the Thames. At the south side of the bridge, you'll find **Butler's Wharf**, a funky section of town boasting multiple restaurants.

3. Get Cultured at the British Museum

Displaying one of the world's finest collections of antiquities, the British Museum contains more than 13 million artifacts from the ancient world. With priceless objects from Assyria, Babylonia, China, Europe, and elsewhere, it's hard to know where to begin in this expansive attraction.

In addition to a well-stocked, on-site bookshop boasting an exhaustive array of titles on ancient history, archaeology and art history, there's a shop selling kids' games and souvenirs, along with one that sells replica sculptures and jewelry.

For those able to linger longer, the museum offers a variety of lectures and workshops, plus a restaurant and café.



4. Get Captivated by Big Ben and the Houses of Parliament



Nothing screams "London" more emphatically than the 318-foot tower housing the giant clock and its resounding bell known as Big Ben. It's as iconic as **Tower Bridge**, and the tolling of Big Ben is known throughout the world. Below it, stretching along the Thames, are

the **Houses of Parliament**, seat of Britain's government for many centuries and once the site of the royal **Westminster Palace** occupied by William the Conqueror.

The best view can be had by crossing Westminster Bridge and looking back. Or, turn left after crossing the bridge and walk along the path towards the **SEA LIFE London Aquarium** (a fun spot to take kids). Gather your crew along the wall for a perfect photo with Big Ben in the background.

Tours of the parliament buildings offer a unique chance to see real-time debates and lively political discussions.

5. Get Creative at the National Gallery



It's almost impossible to visit London without catching a glimpse of the impressive National Gallery. This iconic, columned museum is set at the edge of **Trafalgar Square**, home to incredible masterpieces that make it one of the best attractions in London.

Ranking among the top art museums in the world, London's National Gallery represents an almost complete survey of European painting from 1260 until 1920. The museum's greatest strengths are in its collections of Dutch Masters and Italian Schools of the 15th and 16th centuries.

Among its highlights are a cartoon (preliminary sketch) of the Madonna and Child by Leonardo da Vinci, Michelangelo's *The Entombment*, Botticelli's *Venus and Mars*, van Gogh's *Sunflowers*, and *The Water-Lily Pond* by Monet.

Appendix G

Interpretive phase: Question.

Reading comprehension:

I. Key Word Recognition. Find in the article word/phrase that best expresses the meaning of each of the following words/phrases:

Attractive (1P).....

Addicted to shopping (1P).....

Buy (2P).....

Copy of the original (3P).....

Breathtaking: (1P)

Main Ideas. Write a general idea to the article

Supporting Details. Answer with true, false or not mentioned; then correct the wrong statements; at the end Order these ideas according to their appearance in the text by using a number (1)(2)(3) ...

- a. The houses of the parliament are no more used for political issues, they become a touristic landmark.
- b. The British Museum includes the national treasure and ancient artifacts of the British history.
- c. Big ben was built after a famous politician in London who was responsible for its foundation
- d. The queen and her royal family today take Buckingham Palace as a royal residence.
- e. London attracts a great amount of tourists because of its fascinating landmarks.
- f. London's National Gallery displays one of the world's finest collections from Assyria, Babylonia, China...

IV Organizational Features. How is this text organized?

Choose all that apply and explain briefly why you selected

- B. Narrative
 - C. Descriptive
 - D. prescriptive
 - E. argumentative
- Justification from text:

VI. Inferences. “Read between the lines” to answer the following questions, using information from the article.

- Why the author has written that article?

- According to you, what are her intentions?

I. Author's perspective. Select the perspective or point of view you think the author adopted as he wrote this article and justify your answer with information from the text.

- A. Comic
- B. Factual
- C. Moral/Religious
- D. Persuasive Text

- Justification from text:

VIII. Comparing Cultural Perspectives.

Compare between her love to her country and people in Algeria

Do you think the British love their culture and traditions? How?

What about the Algerian people? Do you think that we still preserving our history; traditions and culture?

Do you think we have such beautiful touristic landmarks that you can describe and convince people to visit?

IX. Personal Reaction to the Text. Using specific information from the text, describe your personal reaction to tourism in London.

Appendix H

Phase One: the Interpretive Scoring Rubric

LITERAL COMPREHENSION

3 points	Key word recognition	Students identify key words in the text that provide clues to the overall meaning of the text.
2 points	Main idea detection	Students identify the main idea(s) of the text.
3 points	Supporting detail detection	Students identify important details that further explain the main idea(s).

INTERPRETIVE COMPREHENSION

2 points	Organizational features	Students identify the ways in which the text is organized and the purpose of organizing their formation in that manner.
2 points	Guessing meaning from context	Students use contextual clues and knowledge of language and text structure to infer the meaning of new words and phrases.
2points	Inferences (Reading Students interpret the overall between the lines)	meanings contained in the text by combining knowledge of key vocabulary, important details, text features, and their own background knowledge.
2 points	Author's perspective	Students identify the author's perspective and provide a justification.
3points	Cultural perspectives	Students identify cultural perspectives/norms by connecting practices/products to cultural perspectives.

Appendix I

Phase two. IPA Interpersonal Mode Rubric with Scoring Points

Language Function ¹

(4points)	Creates with language by combining and recombining known elements; is able to express personal meaning in o basic way. Handles successfully a number of uncomplicated communicative tasks in straightforward social situations, primarily in concrete exchanges and topics necessary for survival in target-language cultures.
(3points)	Uses mostly memorized language with some attempts to create. Handles a limited number of uncomplicated communicative tasks involving topics related to basic personal information and some activities, preferences, and immediate needs.
(2points)	Uses memorized language only, familiar language. Demonstrates narrow range of vocabulary;
(1 point)	Has no real functional ability. Limited range of vocabulary

Text Type²

4 point s	Uses simple sentences and some strings of sentences. thorough development of the topic. very good range of ideas
3point s	Uses some simple sentences and memorized phrases. adequate development of the topic. Adequate d range of ideas, with limited development of the main topic
2 point s	Uses words, phrases, chunks of language, and lists. Limited range of ideas
1 point s	Uses isolated words; fails to provide related support. No range of ideas

Communication Strategies ³

¹ A language task on which the speaker is able to handle in a consistent, comfortable, sustained, and spontaneous manner

² Quantity and organization of language discourse (continuum: word - phrase - sentence - connected sentences – paragraph - extended discourse)

4 points	Responds to direct questions and requests for information. Asks a few appropriate questions, but is primarily reactive. May try to restate in the face of miscommunication.
3 points	Responds to basic direct questions and requests for information. Asks a few formulaic questions but is primarily reactive. May clarify by repeating and/or substituting different words.
2 points	Responds to a limited number of formulaic questions. May use repetition or resort to English.
1 point	Shows hesitation and difficulty to participate in a true conversational exchange.

Comprehensibility⁴

4points	Sentences logically combined	generally understood by those accustomed to interacting with non-natives, although repetition or rephrasing may be required. Valid ideas organized effectively, Variety of appropriate linking devices; very good cohesion.
3points	Sentences fairly logically combined	Is understood with occasional difficulty by those accustomed to interacting with non-natives, although repetition or rephrasing may be required.
2points	Poorly combined	Is understood, although often with difficulty, by those accustomed to interacting with non-natives.
1 points	Many unfinished sentences	Most of what is said may be unintelligible or understood only with repetition.

³ *Quality of engagement and inter- activity; how one participates in the conversation and advances it; strategies for negotiating meaning in the face of breakdown of communication*

⁴ Who can understand this person's language? Can this person be understood only by sympathetic listeners used to interacting with non-natives? Can a native speaker unaccustomed to non-native speech understand this speaker?

Language Control⁵

4 points	Almost no error	Is most accurate when producing effective complex constructions; few errors of agreement, use of appropriate tense. Pronunciation, vocabulary, and syntax are strongly influenced by the native language. Accuracy decreases as language becomes more complex.
3 points	Few minor errors	Is most accurate with memorized language and creating effective but simple constructions, minor problems in complex constructions, some errors of agreement and tense use. Accuracy decreases when creating and trying to express personal meaning.
2 points	Some errors	Accuracy is limited to memorized words. Accuracy may decrease when attempting to communicate beyond the word level. Major problems in complex and even simple constructions, frequent errors of agreement, tenses...etc.
1 points	Many errors	Has little accuracy even with memorized words. No mastery of sentence structure rules, dominance or errors of grammar; lacks basics of sentence structure and variety; meaning is confused or obscured.

⁵ Grammatical accuracy, appropriate vocabulary, degree of fluency

Appendix J

Phase three: the presentational phase scoring Rubric

The total score, obtained from the five evaluated aspects of the language “**Language Function, Text Type, impact, Comprehensibility and Language Control**” as well as “**fluency accuracy and form**” equals 20 points.

1. Language Function⁶: (4 point)

(4points)	Creates with language by combining and recombining known elements; is able to express personal meaning in a basic way. Handles successfully a number of uncomplicated communicative tasks in straightforward social situations, primarily in concrete exchanges and topics necessary for survival in target-language cultures.
(3points)	Uses mostly memorized language with some attempts to create. Handles a limited number of uncomplicated communicative tasks involving topics related to basic personal information and some activities, preferences, and immediate needs.
(2points)	Uses memorized language only, familiar language. Demonstrates narrow range of vocabulary;
(1 point)	Has no real functional ability. Limited range of vocabulary

2. Text Type: ⁷(4 point)

4 points	Uses simple sentences and some strings of sentences. thorough development of the topic. very good range of ideas
3points	Uses some simple sentences and memorized phrases. adequate development of the topic. Adequate range of ideas, with limited development of the main topic
2 points	Uses words, phrases, chunks of language, and lists. Limited range of ideas
1 points	Uses isolated words; fails to provide related support. No range of ideas

3. Impact: ⁸(4 point)

⁶ language tasks the speaker is able to handle in a consistent, comfortable, sustained, and spontaneous manner.

⁷ Quantity and organization of language discourse, on a continuum from words to phrases to sentences to connected sentences to paragraphs to extended discourse.

(4 point)	Extreme Clarity in presenting ideas, well-organized; demonstrates skilful use of cohesive devices, logical sequencing; the writing flows; very fluent expressions, has the ability to maintain the attention and interest of the audience.
(3 point)	Clear ideas, generally organized; some use of some cohesive devices; logical sequenced; limited support logical but incomplete sequencing. Limited ability to maintain the attention and interest of the audience.
(2 point)	Ideas not very clear, Choppy ideas; Loosely organized, clear lack of organizational patterns, very limited ability to attract audience interests.
(1 point)	Ideas disconnected and not clear; follows no clear organizational patter; fails to communicate; no organization or not enough to evaluate.

4. Comprehensibility: ⁹(4 point)

4points	Sentences logically combined	generally understood by those accustomed to interacting with non-natives, although repetition or rephrasing may be required. Valid ideas organized effectively, Variety of appropriate linking devices; very good cohesion.
3points	Sentences fairly logically combined	Is understood with occasional difficulty by those accustomed to interacting with non-natives, although repetition or rephrasing may be required.
2points	Poorly combined	Is understood, although often with difficulty, by those accustomed to interacting with non-natives.
1 points	Many unfinished sentences	Most of what is said may be unintelligible or understood only with repetition.

5. Language Control: ¹⁰(4 point)

4 points	Almost no	Is most accurate when producing effective complex
----------	-----------	---

⁸ *The clarity, organization, and depth of the presentation and the degree to which the presentation maintains the attention and interest of the audience “powerful, memorable, provocative, or moving”.*

⁹ *Who can understand this person’s language? Can this person be understood only by sympathetic listeners used to the speech of language learners? Can a native speaker unaccustomed to non-native speech understand this speaker?*

¹⁰ *Grammatical accuracy, appropriate vocabulary, degree of fluency.*

	error	constructions; few errors of agreement, use of appropriate tense. Pronunciation, vocabulary, and syntax are strongly influenced by the native language. Accuracy decreases as language becomes more complex.
6 points	Few minor errors	Is most accurate with memorized language and creating effective but simple constructions, minor problems in complex constructions, some errors of agreement and tense use. Accuracy decreases when creating and trying to express personal meaning.
2 points	Some errors	Accuracy is limited to memorized words. Accuracy may decrease when attempting to communicate beyond the word level. Major problems in complex and even simple constructions, frequent errors of agreement, tenses...etc.
1 points	Many errors	Has little accuracy even with memorized words. No mastery of sentence structure rules, dominance or errors of grammar; lacks basics of sentence structure and variety; meaning is confused or obscured.

Appendix K

Students' Questionnaire

Dear Student,

You are kindly asked to complete this questionnaire about your Oral Expression Classes during the second semester in comparison with the first semester sessions. Your answers are very important for the validity of our study.

Please, tick the appropriate box (×) to indicate your choice, and specify your answer when needed.

SA: strongly Agree; A: Agree, N: Neutral (or I don't know), D: Disagree , SD: Strongly disagree

Item	Statement	SA	A	N	D	SD
1	Oral expression second semester session is more enjoyable					
2	Assessment in the second semester is not stressful					
3	Oral expression second semester session is more beneficial and informative					
4	Oral expression first semester session is more enjoyable					
5	Oral expression second semester session is so motivating					
6	Reading and writing helped me express my ideas orally					
7	Oral expression session helped me improve my vocabulary					
8	Oral expression sessions helped me develop my ideas					
9	Oral expression teacher cares about enhancing my performance for the better					
10	I do my best to show my teacher and classmates that I learnt from mistakes					

11 Oral expression session helped me develop my communication skills

12 The oral expression session helps me develop all my language skills (reading, writing, speaking and listening)

13 The first activities in oral expression session help to enhance my listening and reading comprehension

14 Reading and listening about the topic helps me to develop my knowledge about the topic I am going to present orally

15 Classroom discussions and comments helps me to adjust my ideas and correct my mistakes

16 Debate and discussions help me develop my ideas and opinion about the topic

17 In the last phase of the session I can rigorously express my final opinion and thoughts about the topic

18 After reading, writing, and listening about the theme I can easily express about it orally

19 My teacher feedback helps me to enhance my performance in all the activities

Appendix L

Examples of Students' Sheets

As the studies showed, The person can get Cold three times a year.

Because being in: well Health is always our priority, try to first avoid the cold by avoiding the close contact with Colleges, washing your hands frequently, and also using the disinfectant gel.

If you got the cold, try to immediately reduce it to not affect your body. Boost your fluid intake of all types but hot herbal teas are the best. try the other relieving ways as the warm gargle with salt, and hot shower.

Honey is also suggested as a natural medicine.

in this period diet is important, start the breakfast with a boost of vitamin C with your fruits for lunch,

take your needed proteins in a bowl of soup to help your immune system. Spices for dinner will be the best choice as it works as antibacterial.

and last but not least, do some exercises to help your emotional wellbeing, and take your rest in full eight hours of sleep.

A general idea of the text:

- The text presented is an explanatory text, titled as "A cure for the common cold", and divided into six paragraphs.
 - Each paragraph contains useful advices and easy to follow instructions on how to act in case of cold infection, effective ways on how to be preventif and strengthes your immunity and your body.
 - Easy and beneficial natural cures are as important as medicinal intake.
 - Not to forget a healthy diet and fluid intake are mostly important to keep yourself well nubrimed and hydrated.
- To beat the flu active reaction you need some good action, exercise and keep your body boosted, and suffisantly rested.

- Reading comprehension: Tourism in London

1) The key words:

a) Attractive: fascinating

b) Addicted to shopping: Shopaholics

c) Buy: purchase

d) copy of the original: Replica

2) The main idea of the article: E) Breathtaking: magnificent, spectacular

London Tourist attraction, and best places to visit in it.

3) True or false statement:

a) false. The houses of the parliament are still used for the political issues, and he give you a unique chance to see it live.

b) false. The British museum includes a lot of treasure and artifacts from all the world, not just British history

c) The big Ben built after the Queen Elizabeth the second (I didn't find it in the article)

D) True.

E) True.

F) ~~false~~ true the finest collections from Assyria, Babylonia, china Europe and elsewhere.

4) The text is organized as: Descriptive

justification: Because the writer is describing the beauty of London for ex C should you hit one of the top museums, or wander around a breathtaking garden)

5) Inferences:

a) She wrote the article to help people to know the most attractive and breathtaking places to visit

b) her intentions is to introduce people to the beauty of her country and attract many tourists.

6) Author's perspective: persuasive text

justification: Because he convinces you with evidence and proofs of beauty of his country (London) and this is by mentioning what is in it (and he is right)

7) comparing cultural perspectives.

a) her love to her country is touching the sky but in Algeria there is some people love their country and there is the opposite

(I didn't understand the question)

b) yes indeed British love their culture and traditions.

d) yes, I think that we still preserving our traditions and culture at least some of them like the traditional food and traditional clothes.

e) yes, we have a lot of beautiful tourist landmarks that can people visit and enjoy their time like Madinet Nasr, the city of thousand domes and the mountains of Djurdjura and the beauty of beaches of Algiers.

- personal reaction to the text:

I have always known that London and England in general one of the most beautiful and wonderful countries in the world and this article gave me more details and more beautiful locations that I did not know about the country.

Tourism in London

Reading comprehension:

I. Key Word Recognition. Find in the article word/phrase that best expresses the meaning of each of the following words/phrases:

Attractive (1P) *impressive - one of the most visited cities.*

Addicted to shopping (1P).....

Buy (2P) *purchase*.....

Copy of the original (3P) *replica*.....

Breathtaking: (1P) *stunning*

Main Ideas. Write a general idea to the article

London is a majestic city and beautiful, luxurious monuments that you have to visit it and you can not miss

Supporting Details. Answer with true, false or not mentioned; then correct the wrong statements; at the end Order these Ideas according to their appearance in the text

a. The houses of the parliament are no more used for political issues, they become a touristic landmark.

False - it offers a unique chance to see real-time debates and lively

b. The British Museum includes the national treasure and ancient artifacts of the British history.

False - the British Museum contains more than 13m artifact from the ancient world

c. Big ben was built after a famous politician in London who was responsible for its foundation

not mentioned

d. The queen and her royal family today take Buckingham Palace as a royal residence.

True

e. London attracts a great amount of tourists because of its fascinating landmarks.

not mentioned

f. London's National Gallery displays one of the world's finest collections from Assyria, Babylonia, China...

False - its the British Museum

IV Organizational Features. How is this text organized?

Choose all that apply and explain briefly why you selected

B. Narrative

C. Descriptive ✓

D. prescriptive

E. argumentative

Justification from text: *because she's describing what's in the country*

.....
Guessing Meaning From Context. Based on this passage, write what the following three words/phrases probably mean

VI. Inferences. "Read between the lines" to answer the following questions, using information from the article.

- Why she has written that article?

..... She wrote this article to give ~~or~~ visitors an idea about London.

- What are her intentions?

..... her intentions was to guide them and to attract visitors to London

I. Author's perspective. Select the perspective or point of view you think the author adopted as he wrote this article and justify your answer with information from the text.

- A. Comic
- B. Factual ✓
- C. Moral/Religious
- D. Persuasive Text

Justification from text: because she's talking about real things about the monuments

VIII. Comparing Cultural Perspectives.

Compare between her love to her country and people in Algeria

..... everyone loves his country in a special way and she does her job. She ~~describes~~ describes with the most beautiful qualities. ppl. in Algeria loves their country but they like to emigrate

Do you think the British love their culture and traditions? How?

..... yes they do, because this country has justice for the poor and the rich they are all the same

What about the Algerian people? Do you think that we still preserving our history; traditions and culture?

..... for the religious holidays yes, but the traditions not too much.

Do you think we have such beautiful touristic landmarks that you can describe and convince people to visit?

..... yes of course our country is so beautiful and charming we just don't know its value.

IX. Personal Reaction to the Text. Using specific information from the text, describe your personal reaction to tourism in London.

..... London is a country that has something for everyone with wonderful cities and luxurious places and I think that its visitors are only high and people who loves ancient things

Tourism in London

- 1) The Answer of Question:
- Attractive: Thousand of visitor
 - Addicted To Shopping = London central, Trendy Sogor
 - Buy = Take coffee
 - Breathing = finest, great
- 2) > True or false =
- a → True
 - b → True
 - c → False, it built after the great fire in 1666.
 - d → True
 - e → True
 - f → Not mentioned.
- 3) > The main idea of the article is =
The tourist destinations in London
- 4) > The Text is organized =
B. Narrative, Because the author talk about the London and her tourist areas. justification from the text is = London is situated in south of England in hours.
- 5) > - She has written that Article, To Attract ~~the~~ tourist London. Show and Tell us the most tourism the oldest Area in London
her attention is = Attract People To visit London

6 > Author's perspective - in Factual
justified: The story of London begins in The Promised Land ...

7 > Comparing cultural perspectives ;
her love to her country is more than her love to people
in Algeria

- yes, The British love their culture and tradition
They are respect their monument and museum and they
don't pollute her. and they are so pride of their culture
and language, they didn't speak any language except
their language.

- yes, I think they are still preserving our history
our tradition, our habit because we didn't change our
tradition.

- yes, in Algeria we have many beautiful places
that people can visit her such as: Timgad,
WADI MIZAB, TIBAZA, EL QANSAH

IX = The personal reaction =

actually, I am very impressed after what I
see and hear about London, it such as a
breathtakingly city, it has many touristic places that we
can visit her, Buckingham Palace, London's eye
Big Ben hour ...

- Short Summary of London :

London the city is the capital of the UK, London is considered a center of economy, transportation and culture. London is located in the South of England, and it is one of the most beautiful tourist and industrial cities in the world as it has the best tourist areas, which they are many, varied and interesting, are considered attractive to tourists, it is one of the most beautiful places in it, which are museums, including the British Museum and the natural history museum, and London is a destination for shopping. Buckingham Palace is one of the most famous ancient buildings in Britain and was the seat of the royal family. The palace is used as a theater for popular events and celebrations in London. and the Tower of London is one of the most amazing buildings. plus Big Ben, London is also there are also beautiful green spaces, many gardens, parks and types of flowers planted for strolling and enjoying the natural places. London is a multicultural city, the city of language and literature. And in the end, London is a beautiful international city name that attracts many tourists to it thanks to its beauty and the beauty of its places.

Résumé

Cette étude visait à examiner l'efficacité de l'évaluation intégrée de la performance (IPA) dans l'évaluation et l'amélioration de la compétence langagière des étudiants. En outre, elle a tenté d'étudier les perceptions et les pratiques d'évaluation actuelles des enseignants universitaires algériens d'anglais langue étrangère (EFL) afin de comprendre les pratiques d'enseignement et d'évaluation de l'expression orale dans le cours sélectionné pour mener l'expérience. De plus, elle a cherché à examiner l'effet rétroactif de l'IPA sur la motivation et l'apprentissage des étudiants EFL. En conséquence, une approche exploratoire séquentielle à méthodes mixtes a été choisie. Premièrement, une méthode descriptive exploratoire a été utilisée à travers l'administration d'un questionnaire en ligne pour étudier les perceptions et les pratiques d'enseignement et d'évaluation actuelles de l'expression orale des enseignants universitaires algériens d'EFL. Deuxièmement, et dans le but de tester l'efficacité de l'utilisation de l'IPA dans l'amélioration de la compétence langagière, un plan pré-test et post-test quasi-expérimental avec groupe témoin a été réalisé. Enfin, un autre questionnaire a été adressé aux étudiants du groupe expérimental. Cette étude a porté sur un échantillon de trente-cinq enseignants de différentes universités algériennes, cinquante-deux étudiants de première année EFL à l'université de M'sila, en plus de trente étudiants pour l'étape pilote. Les résultats ont montré que les enseignants algériens d'EFL ont des attitudes positives envers l'utilisation de méthodes d'évaluation authentiques et basées sur la performance. En outre, il a également été constaté que l'IPA peut contribuer à la promotion de la compétence en matière de compétences linguistiques chez les étudiants EFL. De plus, les résultats ont confirmé la validité et la fiabilité de l'IPA en tant qu'outil d'évaluation. En outre, les résultats ont reflété les attitudes très positives des étudiants envers la mise en œuvre de l'IPA dans leur cours d'expression orale, ce qui montre l'impact positif de l'IPA sur la motivation des étudiants. Enfin, cette étude met en évidence les défis auxquels sont confrontés les enseignants de langues pour intégrer l'IPA et elle évoque plus de recherches dans ce domaine et suggère plus de lacunes à combler.

Mots-clés: évaluation basée sur la performance, évaluation intégrée de la performance (IPA), compétence linguistique, compétences linguistiques, cours d'expression orale.

ملخص البحث

هدفت هذه الدراسة، في البداية، إلى فحص فعالية التقييم الأداء المتكامل (IPA) في تقييم وتعزيز كفاءة الطلاب اللغوية. بالإضافة إلى ذلك، حاولت استكشاف التصورات والممارسات التقييمية الحالية بين أساتذة اللغة الإنجليزية كلغة أجنبية (EFL) في الجامعات الجزائرية في محاولة لفهم ممارسات تدريس وتقييم التعبير الشفوي كونها المقياس المختار لإجراء التجربة. علاوة على ذلك، سعت إلى دراسة الأثر الرجعي للتقييم الأداء المتكامل (IPA) على الطلاب بما فيه تحفيزهم على التعلم.

وبناءً على ذلك، تم اختيار منهج متسلسل تفسيري مختلط. أولاً، تم استخدام طريقة وصفية استكشافية من خلال استعمال استبيان عبر الإنترنت سعياً لاستكشاف التصورات والممارسات التقييمية الحالية لتدريس التعبير الشفوي بين أساتذة اللغة الإنجليزية كلغة أجنبية في الجامعات الجزائرية. ثانياً، وفي محاولة لاختبار فعالية استخدام التقييم الأداء المتكامل (IPA) في تحسين الكفاءة اللغوية، تم إجراء تصميم مجموعة ضابطة قبلية وبعديّة شبه تجريبية. وأخيراً، تم توجيه استبيان آخر إلى طلاب المجموعة التجريبية. شملت هذه الدراسة عينة من خمسة وثلاثين مدرساً من جامعات جزائرية مختلفة، واثنان وخمسون طالباً جامعياً في السنة الأولى من اللغة الإنجليزية كلغة أجنبية في جامعة المسيلة، بالإضافة إلى ثلاثين طالباً للمرحلة التجريبية.

كشفت النتائج أن أساتذة اللغة الإنجليزية كلغة أجنبية في الجزائر لديهم مواقف إيجابية تجاه استخدام أساليب التقييم الحديثة وخاصة المبنية على الأداء. بالإضافة إلى ذلك، تبين أيضاً أن التقييم الأداء المتكامل (IPA) يمكنه أن يساهم في تعزيز كفاءة مهارات اللغة بين طلاب اللغة الإنجليزية كلغة أجنبية. علاوة على ذلك، أكدت النتائج صحة وصدق التقييم الأداء المتكامل (IPA) كأداة تقييمية. إلى جانب ذلك، عكست النتائج المواقف الإيجابية للغاية للطلاب تجاه تنفيذ التقييم الأداء المتكامل (IPA) في مقرر التعبير الشفوي لديهم مما يشير إلى الأثر الإيجابي لهذا التقييم وخاصة على تحفيزهم للتعلم. أخيراً، سلطت هذه الدراسة الضوء على التحديات التي قد يواجهها الأساتذة عند استعمال التقييم الأداء المتكامل (IPA) وأيضاً تدعو إلى التعمق في البحث في هذا المجال وتفتح المزيد من الثغرات لسدها.

الكلمات الرئيسية: التقييم المبني على الأداء، التقييم الأداء المتكامل (IPA)، الكفاءة اللغوية، مهارات اللغة، مقرر التعبير الشفوي.