MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH
UNIVERSITY OF ALGIERS 2- ABOU EL-KACEM SAADALLAH
FACULTY OF FOREIGN LANGUAGES
DEPARTMENT OF ENGLISH

DEVELOPING FIRST YEAR ENGLISH DEGREE STUDENTS' WRITING PERFORMANCE THROUGH ASSESSMENT FOR LEARNING AND THE USE OF RUBRICS AT THE DEPARTMENT OF ENGLISH IN THE UNIVERSITY OF ALGIERS 2

Thesis Submitted in Fulfillment of the Requirements for the Degree of Doctorate in Applied Linguistics and TEFL

Submitted by: Walida Laraba **Supervised by:** Dr. Yasmine Boukhedimi

Board of Examiners

Chair: Prof. Faiza Bensemmane	University of Algiers 2
Supervisor: Dr. Yasmine Boukhedimi	University of Algiers 2
Internal Examiner: Dr. Samira Arar	University of Algiers 2
External Examiner: Dr. Amina Hamdoud	ENS Bouzareah
External Examiner: Dr. Mohamed Chaabane Ali	University of Blida 2

November, 2022

DECLARATION OF ORIGINALITY AND AUTHORSHIP

I hereby declare that the substance of this work is entirely the result of my investigation and that due reference or acknowledgement is made, whenever necessary, to the work of other researchers. I am duly informed that any person practicing plagiarism will be subject to disciplinary sanctions issued by university authorities under the rules and regulations in force.

Date: Thursday, November 9, 2022

Signed: Walida Laraba

DEDICATION

I dedicate this humble doctoral research work to:

the memory of my father, who passed away the first day of this PhD journey. I cannot describe how hard the moments I survived for a long period of time. However, with whom my life was marked by unforgettable and endless support and love, and whose dream was to be what I am today have energetically motivated me to reach this goal, fighting like an injured soldier. My lovely DAD, I never forget you and you are always in my heart and mind. May Allah praise you with Firdaous Paradise

my **mother**, who constantly helps me and believes I deserve the best. May Allah bless you with good health and long happy life.

my **siste**r and two **brothers**, who all love and support me. They are incessantly convinced that I merit higher degrees.

my **husband** who supported me along this long journey with all kinds of assistance.

my primary school teacher**, Medjadba Mahmoud,** who with love and patience inspired us as children with his continuous support, motivation, and care, and who always believed in my capacities to reach higher levels.

all of my lecturers of the different seminars I took in the first two foundational years of this PhD journey, 2016-2017 and 2017-2018, who were very helpful and supportive. I owe tribute and warm thanks to **Dr. Yasmine Boukhedimi**, **Prof. Faiza Bensemmane**, and **Prof. Kamel Khaldi.**

all of my colleagues, **doctoral students of the 2016-2017class**, with whom I spent extremely enjoyable time having fun and debating our queries, with a special dedication to my dear friend and colleague **Abla Benbellal.**

everyone who knows me, near and far, and owes me respect and affection.

you my **dear reader**. I dedicate you this work which may help, inspire, or motivate you.

بِسْمِ اللَّـهِ الرَّحْمَـٰنِ الرَّحِيم

ACKNOWLEDGMENTS

Lasting forever, I would like to acknowledge the help and the assistance of my dear supervisor **Dr. Yasmine Boukhedimi**, who supported me along this long pathway and did never save any effort to provide valuable and constructive feedback. What you have done with me will never be erased. Forever, I am very grateful to you.

I explicitly acknowledge and appreciate the help of **Prof. Dylan Wiliam** from University College London and King’s College London, one of the eminent members of the Assessment Reform Group, who did not save any advice, feedback, source, or guideline to help me carrying out my research.

Also, I highly appreciate the help of **Prof. Royce Sadler** from the University of Griffith, Australia. He was not only an adviser, guide, and helper but also a source of inspiration and motivation. I learned a lot from his experience as he was sharing with me very interesting sources and guidance related to education research. He really boosted my research energy higher.

Particularly, I would like to acknowledge the help of my **participants** of first year at the department of English in the University of Algiers 2 during the academic year 2018-2019, group 10 and 17, for their collaboration and cooperation to answer, respond, and react to all the materials of the intervention with due attention and care. They were very cooperative and kind. Without that, the practical part of this research would never be fully done as so.

The final version of this research work would never come into existence without the **Examination Board**’s feedback. Thank you so much dear examiners for the time we spent together discussing this research piece and the encouragement you awarded me.

Finally, in order to be fair and appreciative **to everyone who helped me** and whose names have escaped my memory, I want to say: I truly appreciate everything you have done for me; whether it was verbal support or providing whatever service I requested.

TABLE OF CONTENTS

Abstract..IX

List of Abbreviations……………...X

List of Tables...XI

List of Figures..XV

List of Appendices...XIX

General Introduction...1

Chapter One: Literature Review 7

Part I. Assessment for Learning: Theoretical and Conceptual Framework……..…..….7

1.1 Assessment for Learning and Formative Assessment…..……………..…………..8

1.1.1 What is Formative and Summative Assessment? ……………..….……………9

1.1.2 Relationship between Formative and Summative Assessment……………...11

1.2 Assessment of and for Learning……………….…………………………………17

1.2.1 Origin of the Phrase ‘assessment for learning’…………………..………....18

1.2.2 Why Assessment for Learning……………………………..……..……...19

1.2.2.1 First Move: Questioning Assessment for Learning and its Relation to Socio-constructivism……………………..………………….…..……...19

1.2.2.2 Second Move: The Relationship between Formative Assessment and Assessment for Learning……………………..…………...………………..21

1.2.2.3 What is Assessment for Learning?...24

1.3 Assessment for Learning and Self-regulated Learning………..………………...26

1.4 Assessment for Learning Implementations ………………………..…………….28

Part II. Classroom-based Assessment: Theoretical and Conceptual framework…..…33

1.5 Feedback…………………………………………..………..……….…....33

1.5.1 Feedback Defined……………………………..………………..……34

1.5.2 Teacher Feedback…………………….……………………………..37

1.5.3 Anonymity in Peer Feedback and Peer Assessment……………..……………40
1.5.4 Feedback and Assessment for Learning…………………….…..…..……….43
1.6 Writing Rubrics……………………..……………………..……….……….……46
1.6.1 What Is Rubric? …………………………………………….…………….....46
1.6.2 Rubrics Classified…………………………………………………………50
1.6.2.1 General vs. Task Specific…………………..……………………………50
1.6.2.2 Analytic vs. Holistic…………..………….....…………………………52
1.6.3 Rubrics' Use…………………...……………..……………….……….…..54
1.6.3.1 Rubrics' Classical Use…………………..……..………………………54
1.6.3.2 Rubrics' Modern Use…………………..…..………………………...55
1.6.3.3 Rubrics in Writing…………………….……………………………...57
1.6.4 Validity and Reliability of Rubrics……………..…………….....………..59
1.7 Self- and Peer Assessment…………………...…..…………………………..60
1.7.1Self-assessment: Definition and Importance….………...….…..………………..60
1.7.2Peer Assessment: Definition and Importance……………….…..…..………66
1.7.3 Relationship between Self- and Peer Assessment…..……………….....……66
1.7.4 Implementing Self- and Peer Assessment………………….…..…....………67
1.7.4.1 Guidelines, Conditions, Procedures and Characteristics…....…………..…67
1.7.4.2 Factors Affecting Self- and Peer Assessment Implementation..…………69
1.7.5 Factors Affecting Self- and Peer Assessment Effectiveness..………..…...…..74
1.7.5.1 Factors Affecting Self-assessment Effectiveness ……………….……….74
1.7.5.2 Factors Affecting Peer Assessment Effectiveness………..…………..…76
1.7.6 Conclusion…………………….……….…..………………………………..78
1.8 Portfolio Writing Assessment ……………………….………….……………..78
1.8.1 What Is Portfolio? …………………………………………………………79
1.8.2 Portfolio for Assessment for Learning……………….……………….……82

1.8.3 Portfolios' Classification……………………………85

1.8.4 Working Portfolio……………………………87

1.8.5 Researching Portfolios……………………………87

1.8.6Portfolio Assessment in Writing……………………………89

Part III. Writing in L2 Context……………………………93

1.9 Approaching L2 Writing……………………………94

1.9.1 Writing Defined……………………………94

1.9.2 Difference between Arabic and English Writing……………………………95

1.9.3 Researching L2 Writing……………………………96

1.9.4 Relationship between L1 and L2 Writing……………………………97

1.9.5 Approaches to Teaching L1 and L2 Writing……………………………99

1.9.5.1 Controlled and Free Composition……………………………99

1.9.5.2 Current-Traditional Approach……………………………100

1.9.5.3 Product Approach……………………………101

1.9.5.4 Writing Process and Process Writing……………………………101

1.9.5.5 Post-Process Approach……………………………105

1.9.6 Writing Difficulties and Challenges Faced by Algerian EFL students ……106

1.10 Paragraph Writing……………………………107

1.10.1 What a Paragraph Is……………………………108

1.10.2 What a Good Paragraph Is……………………………110

Part IV. Empirical Studies……………………………112

1.11 Chapter Summary and Conclusion……………………………119

Chapter Two: Research Design and Methods……………………………121

2.1 Research Questions and Objectives……………………………121

2.2 Context of the study……………………………122

2.2.1 Population……………………………123

2.2.2 Sampling and Sample……………………….……….…………………124

2.2.3 Setting………………………………..……………..…………………125

2.3 Research Methodology…………..…………………………………………125

2.3.1 Study Nature………………………………..………….……………..126
2.3.2 Characterizing the Dependent and Independent Variables ……..……..…...127
2.3.3 Research Design………………..……………………………....………129
2.3.4 Research Methods…………………………..………..…………..………..130
2.3.5 Triangulation…………………..…………………………….…………..131
2.3.6 Validity and Reliability of the Study…………………..……….……….133

2.4 Data gathering tools…………………………………………………….….135

2.4.1 Pre-test and Post-test……………………………………………..……135

2.4.2 Questionnaires……………………………....……………………..….136

2.4.2.1 Designing Questionnaires……………..…………………………....137

2.4.2.2 Validity and Reliability of Questionnaires……………...……...……...138

2.4.2.3 Pre-study Questionnaire……………………………..……..……….138

2.4.2.4 Post-study Questionnaire for the Experiment Group……..……..……141

2.4.2.5 Post-study Questionnaire for the Control Group……………….........142

2.4.3 Writing Rubrics………………..………………………..……….....……143

2.4.3.1 Developing Rubrics for the Current Study………..………..…..….143

2.4.3.2 The Developed Rubrics………………..……………..…….……153

2.4.3.3 Writing Rubrics' Description……………………….....…..………....157

2.4.3.4 Scoring Rubrics…………………..…………………….…..………159

2.4.4 Writing Working Portfolio Project……………………..………..……...160

2.4.5 Researcher Journal Report……………………………...………...……161

2.5 The Study…………………………..……………………………………..…...162

2.5.1 Pre-intervention……………………………………..…………..….…162

2.5.1.1 The Official Writing Syllabus……………………..…………………..162

2.5.1.2 Syllabus of the Intervention …………………………..………………164

2.5.2 Preparing for the Intervention…………………………..……….……....168

2.5.2.1 Discussing Rubrics' Content………………………………….………168

2.5.2.2 Writing Classroom-based Assessment Management……….……….170

2.5.2.3 Familiarizing Participants with Rubrics……………………………..172

2.5.3 The Intervention Proper……………………………………...………..173

2.5.3.1 Pre-requisite Guidelines………………..……………………..……..173

2.5.3.2Assessment Performances………..…………………………...….173

2.6 Data Gathering Procedures………………………..……………………..…174

2.7 Data Analysis Methods and Procedures …………..…………………………176

2.8 Chapter Summary and Conclusion……………………………….…………..179

Chapter Three: Data Analysis and Presentation…………..……….……………180

3.1 Data Analysis and Presentation: First Stage……………………..……..….180

3.1.1 Pre-study Questionnaire Analysis…………..…………………..……180

3.1.1.1 Analysis of the First Section: Background Information……..…..….…181

3.1.1. 2 Analysis of the Second Section: Experience with English Writing and Writing Assessment at Secondary Schools..182

3.1.1.3 Analysis of the Third Section: Attitudes towards English Writing........187

3.1.1.4 Analysis of the Fourth Section: Beliefs about English Writing and Writing Assessment...190

3.1.1.5 Analysis of the Fifth Section: Needs and Awareness towards English Writing 199

3.1.2 Analysis of the Written Pre-test……………………………………......204

3.1.2.1 Substance Errors…………………..………………………………...205

3.1.2.2 Textual Errors…………………..……………..……….………205

3.1.2.3 Discourse Errors………………………..………..………………206

3.1.2.4 Paragraph Structure Problems……………………..………………..206

3.1.2.5 Paragraph Form and Appearance Problems……………..…………..206

3.2 Data Analysis and Presentation: Second Stage………..………..…………208

3.2.1 Analysis of Researcher Journal Report...…208

3.2.1.1 Participant Writing Time……………..…..……………………..………208

3.2.1.2 Time Devoted to Self-assessment Performance…………..……..…….209

3.2.1.3 Time Devoted to Anonymous Peer Assessment Performance………..210

3.2.1.4 Individual Inquiries about Writing Rubric Use………….....……….…211

3.2.1.5 Overall Classroom Assessment……………..………………………...214

3.2.2 Analysis of Participants' Works in the Intervention Proper………..……215

3.2.2.1 Participants with Poor Level…………………………..……………..217

3.2.2.2 Participants with Average Level……………..…..…..……………..224

3.2.2.3 Participants with Good Level……………….………..……………230

3.3 Data Analysis and Presentation: Third Stage……………………………………236

3.3.1 T Test Analysis……………………..………………..……..…………236

3.3.1.1 Parametric Paired Sample T test……………………..……..………236

3.3.1.2. Parametric Independent Sample T test……………..…………..…….237

3.3.1.3 Interpreting T test Analysis………………..……………………....…237

3.3.2 Post-study Questionnaire Analysis………………..…………………..…238

3.3.2.1 Analysis of the First Part of the Post-study Questionnaire…….....……..238

3.3.2.2 Analysis of the Second Part of the Post-study Questionnaire…………..248

3.3.3 Writing Working Portfolio Analysis ………………..……………………255

3.3.3.1 Analysis of the First Part of the Writing Working Portfolio.................255

3.3.3.2 Analysis of the Second Part of the Writing Working Portfolio.............265

3.3.4 Analysis of Questionnaire Administered to the Control Group and Content Analysis of the Post-test... 266

3.3.4.1 Analysis of the Questionnaire Administered to the Control Group......266

3.3.4.2 Content Analysis of the Post-test..270

3.4 Chapter Summary and Conclusion..271

Chapter Four: Discussion of the Findings, Recommendations for Further Research, and Practical Pedagogical Implications ………………………………273

4.1Discussion of the Findings in Relation to the Research Questions…….……..…273

4.1.1 Writing difficulties of first year EFL degree students……………………...274

4.1.2 The Impact of Using Writing Rubrics on First Year EFL Degree Students' Writings…………………………………………………………………...…277

4.1.3 The Impact of Performing Self-assessment on First Year EFL Degree Students' Writings ………………………………………………………...…282

4.1.4 The Impact of Performing Anonymous Peer on First year EFL Degree Students' Writings……………………………………………………...…286

4.1.5 The Impact of Assessment for Learning on First Year EFL Degree Students Writings…………………………………………………………………...…289

4.2 Recommendations for Further Research ………………………………………..294

4.2.1 Further Recommendations for Implementing Writing Rubrics Effectively..294

4.2.2 Further Recommendations for Effective Self-assessment Implementation...296

4.2.3 Further Recommendations for Effective PA Implementation………..……298

4.2.4 Further Recommendations for Implementing Assessment for Learning in Writing Classroom…………..……………………………………………303

4.3 Practical Pedagogical Implications ……………………………………….…304

4.3.1 Pedagogical Implications Regarding the Official Writing Syllabus………..305

4.3.2 Pedagogical Implications Regarding Teachers' Role……….……………..306

4.3.3 Pedagogical Implications Regarding Students' Position in the Classroom...306

4.3.4 Pedagogical Implications Regarding the Writing Staff Members…….……307

4.4 Chapter Summary and Conclusion………………………………………………307

General Conclusion……………………..……………..……………………..308

References……..…………………………...……………………………………311

Appendices..347

ABSTRACT

Assessment for learning has emerged as a means that aims at fostering student learning via basically exploiting assessment to enhance self-regulation, autonomy and lifelong learning. Despite a dearth of evidence in the English Foreign Language writing context, applying assessment for learning is believed to be extremely beneficial. This doctoral research project seeks to raise English Foreign Language students' awareness of a number of core aspects related to writing to improve this productive skill. The core strategy underlying this research study places an emphasis on implementing assessment for learning, which is based on using instructional analytic rubrics as assessment tools, student assessment, including self-assessment and anonymous peer assessment, and teacher guidance. This strategy is used to provide an elaborated and constructive feedback to help students identify their weaknesses and strengths. From the research methodology perspective, this study is applied and deductive, with a quasi-experimental design, and the data collection and analysis methods are blended for triangulation and complementarity purposes. The data was gathered through pre- and post-study questionnaires, pre- and post-tests, writing rubrics and follow-up sections, researcher journal report, and writing working portfolio. The analysis procedures used were various, including content, thematic, and mixed methods and using Microsoft Excel and SPSS version 26. The findings reveal the effectiveness of using assessment for learning to support students' progress and improvement in their writing. The writing of students developed with regards to a number of aspects, namely proper use and acquisition of new vocabulary, correct use of grammar, comprehension and improvement in cohesion and coherence, and writing relevant to the topic. From learner individual perspective, the students admitted to have gained significant motivation for writing as well as increased self-confidence and autonomy, in addition to demonstrating readiness to be lifelong learners. At the conclusion of this study, a set of practical pedagogical implications based on the findings and various suggestions for further research are proposed.

Key Words: Assessment for learning; Writing rubrics; Self-assessment; Anonymous Peer assessment; Teacher scaffolding and assistance; Feedback; EFL writing.

LIST OF ABBREVIATIONS

AfL: Assessment for learning

AoL: Assessment of Learning

APA: Anonymous Peer Assessment

ARG: Assessment Reform Group

EFL: English as a Foreign Language

FA: Formative Assessment

PA: Peer Assessment

SA: Summative Assessment

TA: Teacher Assessment

LIST OF TABLES

Table 1.1 Aspects of Formative Assessment (Black & Wiliam, 2009, p.5) 16

Table 1.2 Synthesis of Main AfL Strategies and Processes (Careless, 2017, p.7) 31

Table 1.3 General vs. Task-Specific Rubrics. From Brookhart (2013, p. 8) 50

Table 1.4 Holistic vs. Analytic Rubrics. From Brookhart (2013, p. 7) 52

Table 1.5 Classification of Assessment Portfolio. From Lam (2018, p.76) 85

Table 1.6 Difference between Controlled and Free Composition by Crookes and Chaudron (1991, p. 52) 100

Table 1.7 Empirical Studies Summary 113

Table 2.1 Participants' Number in Experiment and Control Group 124

Table 2.2 Problems and Solutions for Rubric's Piloting 152

Table 2.3 Activities and Materials Used in Teaching Rubrics 166

Table 2.4 Participants' Choice for the Appropriate Technique 168

Table 2.5 Post-anonymous Peer Assessment Suggestions 171

Table 2.6 Number of Times Required for Using Rubrics and Performing Self-and Anonymous Peer Assessment 172

Table 2.7 Which Analysis Procedure for Which Data Gathering Tool 176

Table 2.8 Content, Thematic and Blended Analysis Methods Used 178

Table 3.1 Errors and Problems Spotted in Participants' Writings 207

Table 3.2 Suggested Assessment Decision Grid 216

Table 3.3 Quantitative/Scoring Assessment Decision 217

Table 3.4 Peer Assessors' Level 217

Table 3.5 Quantitative Assessment and Assessment Decision 218

Table 3.6 Participant’s Reaction to the Follow-up Section in Self-assessment 219

Table 3.7 Participant’s Reaction to the Follow-up Section in Self-assessment in Comparison to the Post-study Questionnaire Results 220

Table 3.8 Peer Assessors’ Level 221

Table 3.9 Quantitative Assessment and Assessment Decision 221

Table 3.10 Participant’s Reaction to the Follow-up Section in Self-assessment 222

Table 3.11 Participant’s Reaction to the Follow-up Section in Self-assessment in Comparison to Post-study Questionnaire Results 223

Table 3.12 Peer Assessors’ Level 224

Table 3.13 Quantitative Assessment and Assessment Decision 224

Table 3.14 Participants’ Reaction to the Follow-up Section in Self-assessment 225

Table 3.15 Participant’s Reaction to the Follow-up Section in Self-assessment in Comparison to Post-study Questionnaire Results 226

Table 3.16 Peer Assessors’ Level 227

Table 3.17 Quantitative Assessment and Assessment Decision 227

Table 3.18 Participant’s Reaction to the Follow-up Section in Self-assessment 229

Table 3.19 Participant’s Reaction to Follow-up Section in Comparison to Post-study Questionnaire Results 229

Table 3.20 Peer Assessors’ Level 230

Table 3.21 Quantitative Assessment and Assessment Decision 231

Table 3.22 Participant’s Reaction to the Follow-up Section in Self-assessment 231

Table 3.23 Participant’s Reaction to Follow-up Section in Self-assessment in

Comparison to Post-study Questionnaire Results 232

Table 3.24 Peer Assessors' Level 233

Table 3.25 Quantitative Assessment and Assessment Decision 234

Table 3.26 Participant's Reaction to the Follow-up Section in Self-assessment 234

Table 3.27 Participant's Reaction to Follow-up Section in Self-assessment in Comparison to Post-study Questionnaire Results 235

Table 3.28 Results of the Parametric Paired Sample T test 237

Table 3.29 Results of the Parametric Independent Sample T Test 237

Table 3.30 Importance of Using Rubrics 239

Table 3.31 Reasons of Using Rubrics if Given a Second Chance 240

Table 3.32 Reasons for Liking Scoring Rubric 240

Table 3.33 Participants' Opinions about Rubrics' Follow-up Section in Self-assessment 245

Table 3.34 Participants' Opinion about Rubrics' Follow-up Section in Anonymous Peer Assessment 245

Table 3.35 Participants' Perception of Writing Classroom Assessment 249

Table 3.36 Participants' Writing Strengths 252

Table 3.37 Participants' Writing between the Beginning and the end of the intervention 254

Table 3.38 Participants' Ordered Works 256

Table 3.39 Writing Practice of the Descriptive and Narrative Genre 258

Table 3.40 Self-assessment Performance of the Descriptive and Narrative Genres 260

Table 3.41 Anonymous Peer Assessment Performance of the Descriptive and Narrative Genres 261

Table 3.42 Rewriting Descriptive and Narrative Genre Writings 263

Table 3.43 Participants’ Perception of Writing Working Portfolio 266

Table 3.44 Participants’ Current Writing Strengths 268

Table 3.45 Participant Writing between the Beginning and the End 269

Table 3.46 Analysis of the Written Post-test 270

Table 4.1 Results Used to Discuss the First Research Question 274

Table 4.2 Results Used to Discuss the Second Research Question 277

Table 4.3 Results Used to Discuss the Third Research Question 282

Table 4.4 Results Used to Discuss the Fourth Research Question 286

Table 4.5 Results Used to Discuss the Fifth Research Question 290

LIST OF FIGURES

Figure 1.1 Assessment for Learning: Its Formative Function and Its Interface with Summative Assessment (Allal, 2011 as cited in Laveault & Allal, 2016, p.8) 23

Figure 1.2 A Framework for Using Portfolios for Learning and Assessment by Klenowski, Askew and Carnell (2006, p. 268) 84

Figure 1.3 Basic Portfolio Characteristics. Extracted from Hamp-Lyons and Condon (2000, p.122) 91

Figure 1.4 Dimensions for Assessing the Portfolio. Extracted from Hamp-Lyons and Codon (2000, p. 144) 92

Figure 2.1 Concurrent Triangulation Design. Adapted from Creswell (2003) 175

Figure 3.1 Participants' Stream in High School 181

Figure 3.2 Writing during Classes in High School 182

Figure 3.3 Presence of Writing Evaluation in High School 182

Figure 3.4 Feedback Given in High School Writing Classes 183

Figure 3.5 Participants' Perception of High School Teachers' Evaluation Method 184

Figure 3.6 Reasons for Liking the Evaluation Method ('Yes') 184

Figure 3.7 Reasons for Disliking the Evaluation Method ('No') 185

Figure 3.8 Reasons for Liking the Evaluation Method ('A little)' 185

Figure 3.9 Self-assessment Techniques in High school 186

Figure 3.10 Peer assessment Techniques in High School 187

Figure 3.11 Whether Liking Writing 187

Figure 3.12 Writing in Free Time 188

Figure 3.13 Writing Preferences 188

Figure 3.14 Writing Anxiety and Stress 189

Figure 3.15 Writing Revision 189

Figure 3.16 Participants’ Perception of their Writing Level 190

Figure 3.17 Participants’ Perception of their Writing Self-confidence 190

Figure 3.18 Justifying Poor Writing 191

Figure 3.19 Justifying ‘Cannot’ Option 191

Figure 3.20 Justifying ‘Do not want to’ Option 192

Figure 3.21 Whether Being a Good User of Grammar 192

Figure 3.22 Overall Grammar Difficulties 193

Figure 3.23 Grammar Difficulties at Sentence Level 193

Figure 3.24 Whether Having Rich Vocabulary 194

Figure 3.25 Sources of Acquiring Vocabulary 194

Figure 3.26 Causes of Having Poor Vocabulary 195

Figure 3.27 Ability to Self-assess 196

Figure 3.28 Explaining Ability to Self-assess 196

Figure 3.29 Explaining Inability to Self-assess 197

Figure 3.30 Ability to Peer assess 197

Figure 3.31 Explaining Ability to Peer Assess 198

Figure 3.32 Explaining Inability to Peer Assess 199

Figure 3.33 Participants’ Ability to Improve their Writing 199

Figure 3.34 Suggested Ways by Participants to Improve Writing 200

Figure 3.35 Face to Face Student-Teacher Interaction 200

Figure 3.36 Reasons for the Need of Student-teacher Interaction 201

Figure 3.37 Reasons for No Need of Student-teacher Interaction 201

Figure 3.38 Whether Preferring Marks over Feedback 202

Figure 3.39 Reasons for Praising Marks over Feedback 202

Figure 3.40 Reasons for Praising Feedback over Marks 203

Figure 3.41 Participants' Suggestion for the Need for Marks, Feedback, or Both 203

Figure 3.42 Participants in Teacher's Position 204

Figure 3.43 Participant Writing Time 209

Figure 3.44 Time Devoted to Self-assessment 210

Figure 3.45 Time Devoted to Anonymous Peer Assessment 211

Figure 3.46 Inquiries' Frequency about the Descriptive Levels 212

Figure 3.47 Inquiries' Frequency about the Follow-up Section in Self-assessment Performance 213

Figure 3.48 Inquiries' Frequency about the Follow-up Section in Anonymous Peer Assessment 214

Figure 3.49 Inquiries' Frequency Overall Classroom Assessment 215

Figure 3.50 Opinions about Using the Scoring Rubric Right from the Beginning in Self-assessment 241

Figure 3.51 Opinions about Using the Scoring Rubric Right from the Beginning in Peer Assessment 242

Figure 3.52 Participants’ Perception of Self-assessment 243

Figure 3.53 Participants’ Perception of Anonymous Peer Assessment 245

Figure 3.54 Perception of Provided Peer Assessment 246

Figure 3.55 Perception of Provided Peer Feedback 247

Figure 3.56 Peer Writings’ Impact 247

Figure 3.57 Participants’ Preference for Anonymous Peer Assessment 248

Figure 3.58 Participants’ Preference for Self-assessment 248

Figure 3.59 Reasons of Involving Participants in Giving Opinions before Doing any Assessment 249

Figure 3.60 Ways of Achieving Progress in Writing 251

Figure 3.61 Ways Helping Participants to Progress 251

Figure 3.62 Participants’ Writing Weaknesses 252

Figure 3.63 Sources of Gaining Self-confidence in Writing 253

Figure 3.64 Ways of Achieving Progress in Writing 267

Figure 3.65 Causes of Non-progress in Writing 267

Figure 3.66 Participants’ Writing Weaknesses 268

Figure 3.67 Sources of Gaining Self-confidence in Writing 269

LIST OF APPENDICES

Appendix 1: Pre-study Questionnaire 348

Appendix 2: Self- and PA of the Descriptive Genre 353

Appendix 3: Self- and PA of the Narrative Genre 354

Appendix 4: Rubrics' Description Levels 355

Appendix 5: Rubrics' Follow-up Sections 360

Appendix 6: Particpants' Memorization Sheet 362

Appendix 7: Involving Participants in Managing Classroom Writing-based Assessment 364

Appendix 8: Researcher Journal Report 367

Appendix 9: Post-study Questionnaire for the Experiment Group (Part I) and Post-Study Questionnaire for the Experiment Group (Part II) 374

Appendix 10: Post-study Questionnaire for the Control Group 382

Appendix 11: Scoring Rubric 384

Appendix 12: Working Portfolio Project 385

Appendix 13: Participants' Writing Progress in the Intervention Group 390

Appendix 14: Analysis Methods Used 390

GENERAL INTRODUCTION

Assessment for learning (AfL), as a concept, emerged in the late 1980s and early 1990s, but it was not widely adopted until 2002 when it was formally widespread by the Assessment Reform Group (ARG) of the University of Cambridge. Since then, it came to explain the shift in the attention given to traditional types of testing from the product of learning to the process (Black & William, 1998). In order to achieve this goal, AfL has been introduced to become an integral component of everyday activities "by students, instructors, and peers that seeks, reflects upon, and responds to information from dialogue, demonstration and observation in ways that enhance ongoing learning" (Klenowski, 2009, p. 264). This is to say that AfL has been emphasized as a means of promoting language progress. Similarly in writing, introducing AfL has been regarded pivotal and not an exception (Lee, 2017), as it may assist students examine their problems, reflect on them while adhering to a set of criteria, and feedback backward and forward to attain pre-defined goals. Thus, this praxis has been strongly recommended for implementation because it is firmly believed to coincide with students' needs for promoting their writing composing strategies.

It is worth bearing in mind that AfL is used sometimes interchangeably with formative assessment (FA) (Harlen, 2005; Bonner, 2013); hence, its principal aim is upgrading students' understanding toward their progress. This is supported by Nelms (2015) who suggests that the progress based on formative feedback can notably improve summative assessment (SA) scores. However, exploiting AfL or/and assessment of learning (AoL) is determined by students' achievement. In this context, Decosta and Reon (2015) assert that "if students are learning less than we hope, then our assessment should lead to more learning. If students are learning as much or more than we had hoped, then our assessment can reveal that happy news" (p.20). We can see that AfL can be reflected in the first 'if part' of this quotation since the situation we want to act upon is 'learning less'. In other words, AfL contributes in engaging students in the learning process, assuming that they may self-monitor and self-regulate their learning, and it can also help them take charge of their own learning (Lee, 2017);

therefore, AfL is expected to help students to be autonomous and lifelong learners. As AfL is seen as a strategy which might improve students' learning, in this study we intend to examine the extent to which it might help first year English as a Foreign Language (EFL), or first year English degree, students overcome their problems in writing. This essentially includes using rubrics, self-assessment and anonymous peer assessment (APA) accompanied with teacher assessment (TA) and scaffolding guidance to build up classroom assessment where feedback is motioning the overall process. In what follows, we explain why this study is worth taking place through highlighting three primordial reasons.

The first reason is related to students' objectives in the degree level. When it comes to teaching EFL writing at the degree level in Algerian universities, the ultimate objective is to help students in producing academic essays that adhere to the writing standards. The journey starts, in the first year at university, by focusing on mastering paragraph writing and understanding the different genres. Therefore, teaching the basics seems an essential step for students to reach their aim. In this context, Grabe and Kaplan (1996) state the need for learning standards as well as paragraph writing for EFL students to advance their English writing to further composing essays and professional articles.

In addition to this reason, many teachers keep complaining about the problems and errors their students make every time despite the attempts made to assist them. Furthermore, students try to not miss the opportunity to express their difficulties in writing and show how much this skill is complex for them. This has been reported in many studies in the Algerian context. For example, Bouyakoub (2012) conducted an exploratory research showing that students had many problems in writing. On the other hand, Khanchali (2017) carried out a study in which he implemented collaborative teaching as a remediating strategy to the low proficiency in academic writing, and he concluded that university EFL students have problems in the writing skill which need urgent interventions.

In addition to the two previous reasons, a third one seems very interesting to shed light on. Most of the time, students feel lost since they are not involved in the process of assessment because teachers often regard assessment as an unwanted task with the potential "to undermine the relationship they have created with their students and the confidence students have gained in their writing" (Hyland, 2003, p. 212). It means that assessment remains an area of strength which teachers intentionally keep students distant from to avoid any interactions or conflicts between them. Being lost might also be as a consequence of students' ignorance of the criteria for revising their writings (White, 1994).

In accordance with what we have discussed above and considering EFL students' writing needs, the current study seeks to achieve a set of aims basically summarized in the following points: (a) Investigating students' difficulties in writing, (b) introducing writing rubrics in addition to student assessment to create an interactive classroom assessment, and (c) assessing the effectiveness of teacher, peer and self-assessment in teaching and learning writing. Because the importance of conducting research makes clear what a particular study would contribute to the literature (Childers & Kent, 2007), the current study has the potential to contribute to improve students' EFL writing in a number of ways. (1) It would provide an in-depth understanding of the difficulties faced by EFL students in writing and help students overcome the challenges via implementing AfL. (2) It could shed light on the importance of students' needs and interests which can enable teachers to know how to satisfy these needs and conduct successful learning. (3) Finally, it would also highlight the role of standardized writing in improving academic achievement of students in other courses. This can prepare them to write coherent and cohesive written samples in exams. According to Gardner (2006), AfL is a process through which students become actively involved in teaching. It notably aims to improve teaching as well as learning (Earl, 2003) and thus it underlines teachers and learners' weaknesses and strengths. Students put onwards a set of goals which they seek to achieve, and teachers play essentially a role of facilitators. When students participate in writing, they may recognize their difficulties and are therefore urged to remedy them through various assessment procedures.

Theoretically, AfL is fundamentally assisted by metacognition in learning, self-regulation, and motivation. First, metacognition, cognition about cognition or thinking about thinking, entangles several processes: (a) recognizing a problem, (b) figuring out the nature of the problem, (c) devising strategies to tackle the problem, (d) monitoring the problem, and (e) evaluating after solving the problem (Sternberg, 1998). In writing, Sternberg's work can be applied accordingly; students highlight their problems in writing, try to recognize the nature of their problems, attempt to solve their problems by developing the appropriate strategies, monitor their writing difficulties by developing their own strategies, evaluate what has been corrected, and feedback on their overall writing process. Second, AfL has lately been articulated by the concept of self-regulation (Clark, 2012), which involves (a) goal setting, (b) self-monitoring with reference to the goal, (c) interpreting and using feedback (e.g., from teacher and peers) arising from self-monitoring, and (d) modification of goal-directed action (e.g., adjusting or redefining the goal) (As cited in Andrade, 2013). These self-regulatory procedures have to go in agreement with the three interests of students: the current goals, the ways of realising them, and the future goals.

Finally, AfL is supported by motivation and this latter focuses basically on learning goals. It stipulates that students are motivated to reach competence rather than performance goals that motivate them toward getting higher scores (Shepard, 2000a). Therefore, AfL can target students' intrinsic motivation; in the classroom, the teacher supervises student learning by treating mistakes as a natural part of learning (Lee, 2017). In this context, the teacher plays the role of a resource, a guide, and a facilitator rather than an evaluator (Shepard, 2000a). In short, AfL enhances learner motivation toward an effective assessment for better learning.

With the study aims and concerns in mind, the following research questions are put forward:

1. What writing difficulties might be highlighted in first year EFL degree students?
2. Does the use of writing rubrics help first year EFL degree students overcome difficulties in writing?

3. Does self-assessment help first year EFL degree students overcome difficulties in writing?
4. Does anonymous peer assessment help first year EFL degree students overcome difficulties in writing?
5. Does assessment for learning help first year EFL degree students improve their writing?

These research questions can be explained through horizontal and vertical rationales. As for the horizontal one, the first research query aims at providing an in-depth understanding of the various problems faced by students in their writings. In other words, it aims at seeking about the writing difficulties. The second research question attempts to examine whether instructional writing rubrics as assessment tools could help students understand and respond appropriately to assessment in the writing classroom, and whether students can make progress. The third and fourth research questions address student assessments, self- and peer assessment (PA), with the aim of examining the effectiveness of these two practices. As for the fifth research question, it can be viewed as a summary of the second, third, and fourth research questions. For the vertical rationale, the first question is designed to identify students' writing problems, as an exploratory stage that requires a second stage where an intervention is needed to remedy to the problems. As a consequence, the subsequent questions are all interrelated to compose AfL and thus come to examine out its effectiveness.

To carry out this study, a quasi-experimental mixed-method design is used to elicit the appropriate data. It is quasi-experimental, or simply experimental in educational research, because it examines the cause-and-effect relationship between variables (Salkind, 2018), which are AfL with its components in relation to EFL writing performance. A mixed method is used to collect data to avoid the pitfalls of each research instrument and to analyze and interpret them (Greene, Caracelli, & Graham**,** 1989). In this study, triangulation and complementarity are behind the use of the mixed method. The former will be required as more than one research method is used and the latter emphasizes a compensatory system, i.e. the weakness of one research method can be adjusted by the strength of the second one (Denzin & Lincoln, 2008).

Moreover, using more than one method can reduce some of the bias that might occur when using only one method (Greene, Caracelli, & Graham**,** 1989; Borkan, 2004; Denzin & Lincoln, 2008; Creswell, 2014).

The data collection tools used to answer the research questions are various. To answer the first research question, a pre- study questionnaire, a pre-test, and a rubric' follow-up section in self-assessment are used to elicit students' writing problems. For the second research question, we relied on the analysis of writing rubrics' content, and their follow-up section of rubrics in self- and APA, researcher journal report, and post-study questionnaire. For the third research question, the pre-study questionnaire, writing rubrics, rubrics' follow-up section in self-assessment, researcher journal report, and post-study questionnaire are used, and the same tools are used to examine APA except for the follow-up section of the rubric. A follow-up section specific for APA is used. Lastly, to answer the fifth research question, a post-study questionnaire, working portfolio, researcher journal report, and pre- and post-test SPSS analysis are all exploited in addition to taking into account to the second, third, and fourth research questions' discussion.

This study is organized into four main chapters in addition to a general introduction and general conclusion including practical pedagogical implications. The literature review chapter examines the main concepts of this study and discusses relevant theoretical frameworks. As for the research design and methods chapter, its content is devoted to roadmap and draw the methodological pathway for conducting the study. After planning research and collecting data, the chapter of data analysis and results will display all the relevant data collected via the data gathering tool mentioned just above. The last chapter, the discussion chapter, is divided into three major sections. The first section is reserved for discussing the findings in relation to the research questions and in relation to the existing literature. The second section presents a set of recommendations for further research generated from research discussion and the last one is devoted to a set of practical pedagogical implications.

1

CHAPTER ONE: LITERATURE REVIEW

This literature review is an embedded and not a stand-alone type as it is an integral part of the study and serves as a context for the topic under investigation (Efron & Ravid, 2019). To this end, the major concepts and constructs related to the objectives pre-set and the formulated research questions are reviewed to delineate appropriate and relevant theoretical as well as conceptual foundations. This chapter, reviewing theoretical and practical works, includes four major parts. The first one is concerned with AfL regarding how it is polished as a distinct concept in the literature in light of other concepts, namely FA, SA, and AoL, while the second part deals with writing classroom assessment where AfL could take place through feedback, rubrics, self- and PA, and working portfolios to flashlight the progress made. In the third part, the main idea is to discuss how to approach EFL writing and the last part is devoted to reviewing the main relevant empirical studies that directly related to the study under investigation.

Part I. Assessment for Learning: Theoretical and Conceptual Framework

The reality of learning can be attributed to Ausubel, who argued over 50 years ago that good teaching begins with putting learners on the progress path instead of just realizing where they need to be. It means knowing what they have learned before deciding upon the next step. Thus, many attempts have since been made to reform the relationship between teaching and learning, with a recent noticeable trend toward exploiting assessment to foster and improve learning, rather than just using it as an evaluative strategy for achievement purposes. Essentially, only through assessment teachers can unveil students’ progress and achievement, and based on the outcomes of that assessment, one can direct a given teaching strategy, make students understand when and why learning occurs, reform a given policy, curriculum or syllabus, and diagnose the psychological learning status of a set of components in students such as motivation, self-confidence, self-esteem, self-efficacy, self-regulated learning, and autonomy.

Exploiting assessment may have originated in the 1960s as a response to Bloom (1968), who claimed that it is critical for instructors to know what learners are learning in order to guide their decisions because learning is unpredictable. This means that the teaching-learning process should be organized in such a way that the desired outcomes can be studied. In this regard, teaching should go in the same direction in order to avoid randomness and to guarantee that the interaction between the two parts, teachers and learners, occurs quietly and either is not chaotic. However, Bloom claimed that it was only under such conditions, it was accepted that students were more successful than others, while the importance of exploiting assessment did not emerge, pointing to the failure of teachers when they reproduced the bell curve of their students' results- i.e. students' grades plotted on a graph called a bell curve, or also known as a Gaussian distribution- rather than fighting what was usually their real job. Recently, exploiting assessment, or AfL, has been demonstrated to have positive effects and any disadvantages stem from how AfL is used, not why it is used (Allal, 2011; Lee, 2017).

It is critical to illuminate AfL's own theory in the literature; nevertheless, it does not have a single, distinct, or even clear theory to be embedded in. It is primarily supported by the contribution of a number of theories, but which exact theories are involved has not yet been determined. For example, the ARG sees that AfL is underpinned by "a theory from the psychology of learning and studies of learning motivation" (ARG, 2002, p.3), whereas Boyd, Green, Hopfenbeck, Stobart (2019), see that AfL draws on three main theories which are: behaviorist theory, constructivist theory, and socio cultural theory. These are the only two clearly stated and unexplained suggestions made by scholars to alert the reader about the lack in the theoretical and conceptual background of AfL. In what follows, we shall discuss the relationship between AfL and FA, AfL and Aol, AfL and self-regulation, and how to implement AfL.

1.1 Assessment for Learning and Formative Assessment

The two concepts ‘assessment for learning’ and ‘formative assessment’ are joined together in the title above because the literature about AfL is surrounded by much of the one about FA as they are inextricably linked. In order to explain the components of the title constructed, including the coordinator ‘and’, we opted for conceptualizing each concept individually and then subsequently examining the relationship between them seeking any interrelationships, intersections, or divergences.

1.1.1 What Is Formative and Summative Assessment?

To begin with, we see that it is important to point to that SA will often appear together with FA because they were used interrelatedly despite the inquiries generated from their relationship. In what follows, we shall basically discuss the origin of formative and SA and the classical view they are seen through.

- **Origin of Formative and Summative Assessment.** It is Scriven (1967) who had coined the terms ‘summative evaluation’ and ‘formative evaluation’ in relation to the evaluation of educational programs and curricula, instructional materials and overall teaching methods. Henceforth, FA and SA were dealt with respectively. Indeed, this idea was further clarified by Bloom, Hastings and Madaus (1971) arguing that Scriven (1967) suggested that summative evaluation is when the curriculum was put in its final form, and formative evaluation took place when a new curriculum was constructed and tried out exploiting the reflections made, before producing the final summative evaluation. Later, Scriven (1996) clarified that both terms belong to one evaluation category.

  From curriculum construction and establishment to the field of student learning, Bloom et al. (1971) borrowed the terms used by Scriven (1967) to explain students’ mastery and improvement of learning through assessment. Despite being drawn from Scriven’s (1967) work, the word evaluation in the formative/summative dichotomy was interpreted by Bloom et al. (1971) in a different way. It is viewed as “the systematic collection of evidence to determine whether in fact certain changes are taking place in the learners as well as determine the amount or degree of change in individual students” (p.8). This is to say that it reflects the cursory progress and any change made across a learning continuum to assist learners achieve their final learning targets.

- **Classical View of Formative and Summative Assessment.** SA is traditionally associated with judging, grading, and certifying student achievements at the end of a course or a program, while FA tends to help in the teaching and learning process by allowing teachers to intervene when students are learning and adjust their instructions accordingly (Bloom et al., 1971). On the contrary, introducing FA at the end of each unit is thought to be useful for providing feedback on what students have learned and what they need to learn better, via introducing corrective activities such as additional resources and auxiliary tasks (Bloom, 1968) and varying the teaching instructions to go in line with students’ learning objectives (Bloom et al., 1971). The clear distinction between summative and FA, in fact, was provided by Scriven (1991). This author preferred to refer to what he called Stake’s Maxim (Robert Stake) as follows, “perhaps the best way to put the formative/summative distinction is due to Robert Stake: when the cook tastes the soup, that’s formative evaluation; when the guests taste the soup, that’s summative evaluation” (p. 19). The guest is the teacher and the chef is the student. For a better understanding, we examined the link between FA and SA in the forthcoming Section 1.1.2.

The idea of FA proposed by Bloom (1968) and Bloom et al. (1971) lies in modifying each time the teaching process based on the assessment learning outcomes to verify whether the measures used helped students to learn. Guskey (2005), on the other hand, suggests that the second FA, which represents the outcomes of the first FA, could motivate students through giving them another chance to succeed; however, this could only happen if FA instructions are aligned to the overall course objectives to help students toward better achievement in SA (Bloom, 1976). Following Bloom’s (1968, 1976) writings, FA has three major roles: (a) providing feedback to students to inform them about their current position in learning and how they may advance to a better one, (b) offering feedback to educators to revise and adjust their instructions, and

(c) providing students with opportunities to be motivated and to enhance their learning.

Based on the literature, FA is seen to be more associated with traditional didactic teaching methodologies and concepts as some scholars suggest. For example, Torrance (1993) perceives Bloom’s views on FA as mechanistic and sees that it would function efficiently when there is an interaction between educators and learner, as it is recognized in the socio-constructivism orientation. This kind of criticisms toward FA and SA, namely being mechanistic, is conceivable because no theories were developed to shape both assessments, leaving the work unsubstantiated and subject to different unjustified objections. It is important to recall that the discussion regarding FA and SA is not our primary interest; therefore, we have opted to focus only on the major conflicts which marked this research field. All in all, we can view FA as a precursor to SA achievement, and both interplay and complement each other in one way or another, regardless of being theoretically framed or not. Further understanding is supplemented by the forthcoming Sections (1.1.2 & 1.2).

1.1.2 Relationship between Formative and Summative Assessment

After being adopted by Bloom, many researchers such as Black and Wiliam (2003) and Guskey (2005) reported that many educators attempted to implement FA in their classrooms in different places around the world. However, this concluded the mismatches between theories and practice. Biggs (1998) is the first to re-establish the relationship between FA and SA. He criticized the eminent review made by Black and Wiliam (1998) and pointed to the exclusion of SA and its effect on learning. As a response to that, Biggs (1998) and Taras (2005) consider both assessments important for learning. Despite being born together and emphasized implicitly from the beginning (Scriven, 1967) and explicitly thereafter (Bloom et al., 1971; Scriven, 1991), the relationship between FA and SA witnessed fluctuations, essentially due to (a) external pressures for certification and accountability, (b) confusions made about the definition of FA, and (c) the lack of understanding of the concept of FA including the lack, not the absence, of the theoretical underpinnings. These reasons are fully discussed below, with further elaboration of the third reason due to its complexity.

a.External Pressure for Certification and Accountability. External pressure for certification and accountability has altered the relationship between FA and SA. In higher education, for instance, Elton and Johnson (2002) assert that assessment is “still pervaded by a largely unreflective traditionalism” relying on assessing students summatively via examinations, essays or reports (p. 9). Besides the attainment standards, other external constraints have been highlighted by Yorke (2003) such as the increasing student/staff ratios and the difficulties generated by the measures of implementing FA, namely being disciplined and fostering teachers to act as active researchers in classroom to reflect upon the curricula. Furthermore, Atkins, Beattie and Dockerell (1993), in a less sympathetic view, see that teachers in higher education are still clinging to conservative assessment as a result of “ignorance or unwillingness to consider change” (p.26). Consequently, FA implementation and exploration were delayed.

b.Confusions Made about the Definition of Formative Assessment. The Poor understanding is also the result of the confusions over the definition of FA as Yorke (2003) referred to as ‘definitional fuzziness’. This fuzziness was first explained by Sadler (1989), by pointing to the minor importance given to feedback and FA which makes it then “mostly hortatory, recipe-like and atheoretic” (p.122). In this context, Torrance (1993), for example, stated that FA was still seen fragmented and not well explained. By tracking that blur, we found it still exists. For example, Bennett (2011) asserts that the term FA does not yet “represent a well-defined set of artefacts of practices” (p. 5). Therefore, FA’s misunderstandings has led to ignorance and disregard, either intentionally because of the difficulties encountered in implementation or unintentionally due to misunderstanding the concept itself and its aim..

It is clearly stated that FA definitions remain contradictory among scholars and researchers. The problems stem from issues such as whether or not grading should be considered, whether or not it should be a continuous process, and whether it should be formal or informal. In what follows, we will briefly explain how these conflicting views in the definitions of FA have been raised. Sadler (1989), the first who prompted fuzziness in FA, suggests that awarding marks has a counterproductive effect and FA should be exempt from giving them, while Miller, Imrie and Cox (1998) see that mark can be assigned because they have a neglected effect on students' final results. We can explain this into two ways: (a) including marks given regularly in the final grade might jeopardize validity of SA, or (b) being severely addicted to marks might prevent students from making progress. Another conflicting view is continuity in FA. Brown (1999), for instance, argues that FA must be continuous to be effective but Yorke (2003) sees that the effect has nothing to do with continuity as long as the ultimate goal is met. So it can be occasional. The final point raised is informality and formality along a continuum. Along a continuum of formal and informal assessment, Rowntree (1987) proposes that formal FA takes place when it happens with curricular assessment framework, whereas the informal one occurs in the course of event without being explicitly prescribed in the curriculum design.

As its definitions are poorly constructed, understood, and formulated, FA's components and constituents, as a consequence, are also inadequately perceived both in school (Black & Wiliam, 1998) and higher education level (Shavelson, Black, Wiliam, & Coffey, 2004). In a nutshell, FA has faced difficulties in practice which have led scholars to redefine it each time, hoping to simplify its practical instructions and guidelines. This has created further discrepancies between theory and practice and has congested the literature with many definitions, which are difficult to synthesize or even conceptualize for further implementations (Bennett, Wragg, Carre, & Carter, 1992; Russell, Qualter, & Mcguigan, 1995; Yorke, 2003).

c.Lack of Theoretical Underpinnings of Formative Assessment. The poor understanding could also be explained by the lack, not the absence, of the theoretical underpinnings of FA, despite a number of major attempts made by Sadler (1989), Yorke (2003) and Black and Wiliam (2009). Sadler's (1989) work, the first who attempts to formulate a theory of FA, came fundamentally as a response to the behaviorist usage of feedback and FA. He argued that the cyclic implementation of formative feedback- the outcomes of the first implementation contribute to the improvement of the second, and so on- is only valid in simple learning, but not in the complex one. Simple learning is supposed to occur only when to say that is right or wrong and where feedback stimuli would work in some subjects and not in all. However, complex learning needs judging student learning qualitatively rather than as being just right or wrong. Toward this aim, FA and feedback should be different when feedback is more than knowledge or results in order to guide students into “a concept of quality roughly similar to that held by the teacher”, and for students to be able “to monitor continuously the quality of what is being produced during the act of production itself, and has a repertoire of alternative moves or strategies from which to draw at any given point” (Sadler, 1989, p.121). From this, we can understand that FA and feedback should act together in a way to help students take charge of their own learning and be able to recognize their learning objectives through exploiting all the assessment outcomes toward better improvement and achievement.

To frame his theory, Sadler (1989) proposes three conditions that must be met simultaneously, focusing fundamentally and intentionally or unintentionally on the learner because this latter has to (a) “possess a concept of the standard (or goal, or reference level) being aimed for”, (b) “compare the actual (or current) level of performance with the standard”, and (c) “engage in appropriate action which leads to some closure of the gap” (p.121). To explain these conditions in relation to the reality in classrooms, we find that the first condition requires not only students but also teachers to be aware of the implicit and explicit standards, both together, and to understand their goals in order to find ways to achieve that. All of this must form a class unit, which is primarily characterized by the interaction of the parts involved there. To meet the second condition, students must be able to objectively assess their own work against the criteria. Therefore, teachers’ role should not only be limited to providing those criteria but also helping them understand why and when to practically use them. Finally, to encourage students narrowing the gap between the initial and the target levels, students should interact with each other and with the teachers to bring their understanding to the final target.

Unfortunately, the author failed to formulate a theory for FA as he did not establish clearly the underpinnings of his theory because he stated explicitly that it is framed in Deweyan and Vygotskyan works. This prevented his theory from being independent with well-defined dimensions. Another ambiguity in this theory is the repositioning of teaching and learning. This is seen as both difficult and broad at the same time, and the halo surrounds the idea of repositioning itself. This questions what to reposition: the curriculum, the teaching and learning policy, the teacher, or the learner. Finally, another problem that contributed in this ambiguity is discarding the summative component even though it is emphasized to be an integral part that always appears with FA. In fact, the importance is not to formulate a clear theory of FA, but rather to avoid the discrepancies that might be generated by these outlined conditions and their relation to the practice when being introduced and activated.

Another less important attempt to develop the theory of FA was made by Yorke (2003). Toward this aim, the author suggested considering the following principles:

- the epistemological structure of the relevant subject,
- the students' ontology including both psychopathology and development,
- relevant theoretical constructs related to learning and assessment,
- educator/assessor's professional knowledge (not only knowledge about the discipline but also about student development at every level in addition to knowledge about assessment methodology and the psychology related to feedback delivery/giving/receiving), and
- relevant theory for communication and interpretation.

Based on the above principles, this attempt was broad and non-specific and not embedded in well-defined theoretical foundations. For further clarification, the same author explains that the constructs he outlined to underpin FA are universally appreciated by university teachers, as in higher education they focus on their own specific subject disciplines. This would be true as teachers in higher education should possess a high level of critical thinking and should not practice teaching in a systematic and robotic manner.

The third attempt to develop FA theory was made by Black and Wiliam (2009). Despite stating clearly that they "didn't start from any pre-defined theoretical base" (p.5), the authors seem following the same path as Sadler (1989). And the only difference which is clearly noticeable is the definition of the classroom dimensions. This latter focuses on the parts involved and how those parts should react in each of the three learning stages. The following table 1.1 displays the three stages of learning with the corresponding interventions of the parts involved.

Table 1.1 *Aspects of Formative Assessment (Black & Wiliam, 2009, p.5)*

	Where the learner is going	**Where the learner is right now**	**How to get there**
Teacher	Clarifying learning intentions and criteria for success	Engineering effective classroom discussions and other learning tasks that elicit evidence of student understanding	Providing feedback that moves learners forward
Peer	Understanding and sharing learning intentions and criteria for success	Activating students as instructional resources for one another	
Learner	Understanding learning intentions and criteria for success	Activating students as the owners of their own learning	

In light of Table 1.1, the first pitfall spotted is that the authors have only listed the concepts underlying FA without emphasizing the actual implementation. The components (teacher, learner, and peer) and the phases (the current level, the target one, and how to best move from the initial to the target level) should not work in a linear way but they should be more interactive, somewhere at least. They assigned each member a specific role(s) to fulfill at a specific level of the gap-reducing metaphor while overlooking any relation between them. They see that this suggestion is based on "moments of contingency" (p. 8) and the inclusion of the peer and the learner is like in the Vygotskyan and Deweyan social learning. Similarly to Yorke's (2003) work, they stressed the need for theories on instructional design, curriculum, pedagogy, psychology and epistemology. In fact, Black and Wiliam (2009) identified various models and examples throughout their paper but all resembled Sadler (1989) and Yorke's (2003) works, who all fail to establish clear theoretical underpinnings for FA.

To conclude the three attempts that have been made to theorize FA, we would say that the authors failed to formulate a theory of FA as this might be the result of the confusion caused by the large number of definitions established for this concept, which were all not fully flowing in the same direction more than generating contradictions. That might also be due to the effect exerted by the different theories underpinning learning, ranging from behaviorism to socio-constructivism. The former stance de facto could strongly affect the decision to be made; either to consider all theories together which might be quite impossible to happen, or to prefer one over the other which again was automatically rejected, as the three trials turned out to be.

To end the discussion regarding the relationship between FA and SA, we would say that the poor understanding and lack of well-defined theoretical foundations of FA are the main reasons behind the development of formative/ summative assessment dichotomy. This is viewed as if there is a rivalry between the two concepts and not a kind of complementarity. Recently, Wiliam (2020) reconsidered the idea by suggesting that the same results obtained from a given assessment can be used formatively and summatively; therefore, the words formative or summative would be better seen as two different conclusions rather than two different kinds of assessments. In line with this, Barnet (2007) has already argued that "summative assessment is itself formative" (p.36). The same author has focused more on whether the whole, FA and SA, can conclude to positive outcomes, qualities, and dispositions in students rather than delving deeper in either type alone and discarding the second component. One would wonder whether the cut between FA and SA has always generated troubles among scholars and, more severely, one would question what mostly have led to unsatisfying results. In line with this, we could also discern another major problem disrupting the relationship between FA and SA, namely the attempts to develop theories about FA. This led to problems arising from trying to apply the theoretical considerations related to both assessments. Taking all together and regardless of developing a theory of FA, the main idea that should be at the forefront is to encourage students to be active learners and to develop reflective thinking through assessment and feedback, without breaking the relationship between FA and SA and by incorporating all the elements of the classroom into one interactive unit.

1.2 Assessment of and for Learning

Before embarking on the discussion about how assessment should take place as a new orientation in learning and teaching process, it seems convenient to start with summarizing the importance of learning in the following quotation by Cameron, Tate, Macnaughton, and Politano (1998). These authors state

> learning occurs when students are thinking, problem-solving, constructing, transforming, investigating, creating, analysing, making choices, organising, deciding, explaining, talking and communicating, sharing, representing, predicting, interpreting, assessing, reflecting, taking responsibility, exploring, asking, answering, recording, gaining new knowledge, and applying that knowledge to new situations. (p. 6)

To support learning and help students fulfill what the aforementioned scholars have termed learning objectives, recently many researchers proposed exploiting assessment. For example, Taras (2007a) claims that the proposed idea which is assessment should promote and support learning “brooks no denial” (p.58). This is reflected in Dewey’s (1933) ‘learning loop’, Lewin’s (1952) ‘reflective spiral’, Schön’s (1983) ‘reflective practitioner’, Senge’s (1990) ‘reflective feedback’, and Wiggin’s(1993) ‘feedback loop’. To explain assessment relationship to and with learning, we will discuss how assessment would be used to inform learning for better help, improvement, progress, and achievement i.e. AfL.

1.2.1 Origin of the Phrase ‘Assessment for learning’

AfL is a term that has significantly surfaced the scholarship of learning and teaching over the last decade. This term was generated from the eminent review made by Black and Wiliam (1998) on the positive effects of FA on students’ learning. Before defining this concept, the first step that merits to start with is clarifying some of the confusions made about the origin of this phrase. The phrase ‘assessment for learning’, as its meaning suggests, first appeared in a paper by James (1992) at the annual conference of the Association for supervision and Curriculum Development in New Orleans. Five years later, this phrase was used by Gipps and Stobart (1997) in the second chapter of the third edition of the book entitled *Assessment: A teacher’s Guide to the Issues.* Despite the fact that it has been mistakenly attributed to Stiggins (2002), which he himself denied and attributed it to the UK authors, this term was brought to a wider audience by the ARG (2002) of Cambridge University in a brochure entitled *Assessment for learning: 10 principles. Research –based principles to guide classroom practice*.

1.2.2 Why Assessment for Learning

The assessment debate is “awash in hidden assumptions, unstated values, partial truth, and confusions of ideas, false distinctions and irrelevant emphases” (Rowntree, 1987, p. 3-4). More than two decades of Rowntree’s testimony, it appears to be true that in the literature assessment debates have involved hidden assumptions and many new terminologies have emerged and may continue to emerge. The latest interest is using assessment to promote learning, and this has so far been supported to be AfL. The use of AfL is marked by two main transitions; the first is from AoL to AfL and the second from formative to AfL.

1.2.2.1 First Move: Questioning Assessment for Learning and its Relation to Socio-constructivism. Moving from AoL to AfL is emphasized by many authors. Some eminent scholars can be listed such as Black and Wiliam (1998), Torrance and Pryor (1998), Gipps (1999), and Shepard (2000). This move came as a response to two main reasons.

- **First Reason.** It is the negative effects of AoL that are noticed on students’ learning outcomes in particular being score-dependent and focusing on quantity rather than quality. This resulted from the uncontrolled importance given to AoL in learning. In higher education, for example, Knight (2002) has compared that importance to ‘Achilles’ s heel’; meaning that AoL is the pillar that makes learning outstands, otherwise learning would have no value or existence. To enrich the discussion and for further understanding, we shall attempt to define AoL.

AoL is viewed differently. It is found in the literature to be directed against AfL and is not clearly stated to be equated with SA in the works of some scholars, namely ARG (2002), Stiggins (2002), and Wiliam (2007). It also seems to be opposed to AfL and clearly equated with SA in works such as the one by Earl (2003). In fact, AfL vs. AoL distinction did not attain a final clear stance but rather various and contradictory views. Recently, AoL has been seen complementary to AfL and FA (Allal, 2011). In the same context, Laveault and Allal (2016) see that AoL can start AfL mission; spotting some aspects in learning to work on via adopting AfL. This can be done via, for example, using diagnostic tests to account for the initial level of students in terms of knowledge, learning capacities, and/or skills before embarking on learning proper.

Regarding the distinctions made, AoL is also seen; differently, but essentially for one purpose which is measuring learning. It is used for grading and measuring purposes (ARG, 2002; Laveault, & Allal, 2016), and measuring learning “after the fact” (Earl, 2003, p.25). Unlike AfL, AoL differs from assessment, which is primarily intended “to serve the purposes of accountability, or of ranking, or of certifying competence” (Black, Harrison, Lee, Marshall, & Wiliam, 2003, p. 10). Regardless of being or not equated to SA, opposed to AfL or not, or/and possibly having synergic functions with AfL, and having various definitions of AoL, this concept can be viewed as a step in learning which measures it. It often comes at the end, but it can also come at the beginning or even in the middle of the learning process.

- **Second Reason.** The second reason underlying this move goes in agreement with the change in the meaning of learning, shifting from passive to active knowledge within socio-constructivism which leads to changing the world view and epistemology (Gipps, 1999). This change should also be accompanied with the change in assessment since learning and assessment are inextricably interrelated (Gipps, 1999; Elton & Johnston, 2002). Thus, if the teaching method is changed but the assessment is not, the result probably concludes to an automatic failure (Elton & Johnston, 2002). In relation to constructivist learning, Adams (2006) suggests that the latter requires accepting individual differences in building knowledge through various ways like selecting, acquiring, interpreting, and organizing information. As a result, opting for various assessment strategies, within AfL context, to foster learning should be undeniable to align the learning objectives to the instructions established. In fact, many constructivist theories are found in the literature all of which have a common conceptualization which is learning is an experience that is contextualized for the sake of dialoguing, discussing, and solving problems (Adams, 2006). To achieve this, AfL could create a community of practice where students as well as teachers could shape their knowledge about learning and teaching through discussion and critical interactivity. This should be encouraged by student-student and student-teacher exchanges in different directions; interaction of one-to-one, peers, group, and from student to teacher and in the other way round.

1.2.2.2 Second Move: The Relationship between Formative Assessment and Assessment for Learning. The relationship between AfL and FA is found multidirectional, specific, and determined. As noted, FA came first through Scriven's (1967) work while AfL came second after the seminal work by Black and Wiliam (1998) where more than 250 works about FA were reviewed. After analyzing the works exploring and examining the relationship between FA and AfL, we have decided to divide the critics into three main groups: (a) the proponents of equality between FA and AfL, (b) the inequality between the two concepts, and (c) other miscellaneous critics about AfL.

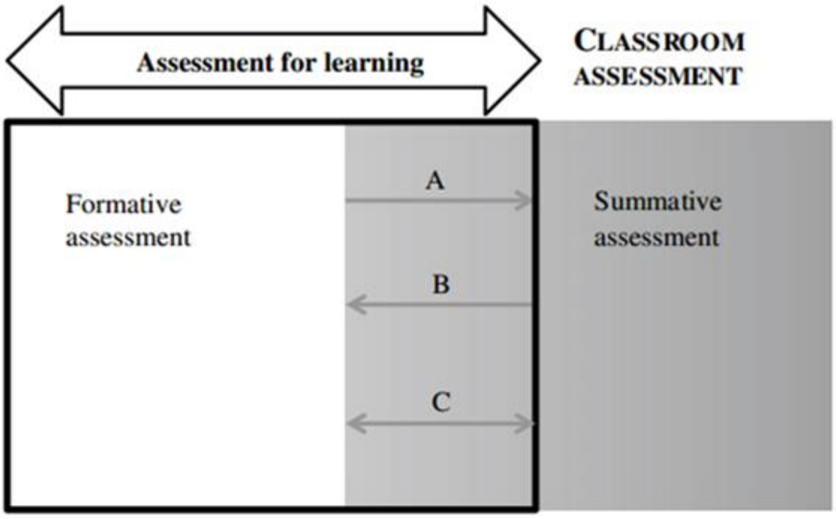
a.Assessment for Learning is equated to Formative Assessment. In this category, some scholars foster the equity between the two concepts, FA and AfL. For example, Stiggins (2009), Bennett (2011), and Calkins, Ehrenworth, and Akhmedjanova (2019) see that AfL, or FA, is used to help students learn and inform the instructional decisions. This equality is featured across different dates until recently, which implies that this idea still exists and is valid in the eyes of many scholars. However, Bennett (2011) explains that the change encourages shifting the definitional burden.

b.Assessment for Learning is not equated with Formative Assessment. Moving from FA to AfL, or the non-equity between the two, is supported by many scholars. For example, in an interview with Professor Dylan Wiliam, one of the eminent members of the ARG, conducted by Bethan (2002), Wiliam comments that he "mourn[s] the loss of the term 'formative' ", as he feels that "the formative dimension, the requirement to form the direction of future learning, places a slightly stronger imperative on a teacher to really make it count whereas assessment for learning can actually sound like a prescription to assess in order for the student to learn" (p.48). To explain this, Stiggins (2002) posits that AfL is not FA despite the attempts to achieve equity as the latter is more with "testing frequently or providing teachers with evidence" and its outcomes are used to revise the instructions whereas the former "must involve students in the process" in classroom assessment to check its understanding and uses of "the continuous flow of information about student achievement that it provides in order to advance" (p. 5). As such, AfL focuses on involving students to make informed decisions in both teaching and learning. In this vein, Black and Wiliam (1998) acknowledge the importance of FA; nevertheless, they emphasize that using assessment to improve learning should be AfL as the former, FA, focuses on the function and the latter, AfL, on the purpose.

More than 20 years later, Wiliam (2020) still supports the same idea, arguing that AfL is meant to improve learning and not measure it, and it is more than just relying on feedback as FA does. In criticizing FA function, Stiggins and Chappuis (2006) prefer the term AfL over FA as they feel that this latter is becoming very narrow because it seems linked to "more frequent summative assessments administered at regular intervals" (p.1). In addition to this explanation, Wiliam (2007) addressed the phrase 'assessment for learning' clarifying that it should not be equated with measuring student learning or with AoL as it is more about supporting learning, interactivity in the classroom, and bridging between teaching and learning. Thus, the main idea about the impact of AfL is twofold: keeping students engaged in learning and remaining confident to progress if they keep trying to learn (Stiggins, 2002). In other words, AfL does not force students to learn but gives them opportunities to learn. In agreement with this, Stiggins (2002) furthered saying, “in short, everyone wins. There are no losers” (p. 9). He means that all the parts involved in AfL gain as it pinpoints learning as a multifaceted process, regarding teaching instructions and principles which might be shaped accordingly.

In an attempt to determine the relationship of each concept to the other, formative and AfL, Allal (2011) suggests that FA is just a part of AfL. She illustrated that in Figure 1.1.

Figure 1.1 *Assessment for Learning: Its Formative Function and Its Interface with Summative Assessment (Allal, 2011, as cited in Laveault & Allal, 2016, p.8)*



A: information from FA is taken into account in determining a summative assessment.

B: information from a summative assessment is used in a formative manner to support learning.

C: an assessment activity is composed of phrases or components some of which have a formative function and others a summative function.

Figure 1.1 proposed by Allal (2011) shows the components of classroom assessment carried out under AfL. This encompasses FA, SA and the intersection between FA and SA. The intersection zone is represented by three capital letters: (A), (B), and (C). This zone is created when SA is used to inform FA and vice versa and when exploiting activities or any materials, which they require the presence of both assessments. This means that FA and SA are overlapping notwithstanding the functions they fulfill. Briefly, AfL for this group of scholars means SA, FA and the interrelation between the two to form classroom assessment

c.Assessment for Learning is explained in different ways, other than the two former ones. More than being equated to or different from FA, AfL have been criticized by the third group of scholars, though few in number, in various ways. For example, Boud and Falchikov (2007) urged teachers to determine the factors which limit assessment to improve learning, but they ignored the long term goals higher education works for. On the other hand, Careless (2007) instead proposed 'learning-oriented assessment' as an alternative to AfL especially when assessment is used to improve learning, and he did not originally accept the appellation, AfL. Another explanation is suggested by Good (2011). The author recommended the use of the expression 'formative use of assessment information' (p. 1) to express the idea of exploiting assessment information to determine its function, and Earl (2003) instead proposed 'assessment as learning'. The meaning of this phrase denotes formative AfL; it is seen as a feature which might emphasize the role of students as having critical roles in assessment and learning and not solely as passive receptors. To be clear as to these explanations and suggestions of this category and up until this moment, it seems worth highlighting that this kind of comments about AfL has no influence on the whole literature about AfL because they have not been thoroughly discussed and they have remained as truncated propositions because they are not fully documented or given importance. This is simply because AfL is inherently controversial, and opening up other debates could make AfL discussions worse and even weaker.

To close the discussion, we have noticed, from our vantage point, that scholars are trying to bring AfL to the horizon instead of causing its loss in a dark ambiguity as happened with FA. None of the scholars can be blamed; on the contrary, their discussions were very constructive and enriching as they have unveiled much confusion. However, the ambiguity we mean here is the inability to come up with one/ or a set of agreed upon definition (s) of AfL, at least, for further effective implementation. Moreover, we did not, either un/intentionally, mean that FA was left completely in the corner but we meant that the scholars' attempts concerning AfL are much more focused on clarifying and making it understandable to foster its implementation, overlooking any confusion that might prevent that. That is why scholars are currently not interested in generating new concepts and discussing them, such as the ones suggested by the scholars of the third category, but only being interested in AfL and FA and the interrelated assessments, namely SA and AoL.

1.2.2.3 What is Assessment for Learning? Before we start defining AfL, we want to clarify that the definitions put forward are proposed by scholars who foster non-equity between AfL and FA and those who implicitly explains that FA is AfL. To clarify the last point, 'those who implicitly explain that FA is AfL', recently many scholars explain FA as AfL, digressing the focus from the function to the purpose without explicitly stating it as AfL. It is worth pointing to that the very recent literature about FA, most of it, is implicitly exceeding the boundaries to AfL, but without declaring that explicitly. It means that many scholars express that this is FA but their discussion is purely flowing into AfL. This might be caused by the accumulated confusions existing in the literature and/or misunderstanding of the conceptualization of the concepts.

Despite being highly emphasized and accepted to strongly and positively impact learning and having a high influence in attracting policymakers, practitioners, and researchers, the AfL definitions have not yet been fully formulated and developed. To explain this, Murphy (2006) sees that AfL is still described as "a neat catchphrase" that needs further definitions (p.42), and Careless (2017) adds that AfL "is often not defined explicitly" (p.4). The first definition worth starting with is the one suggested by the ARG. They defined AfL as "the process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how best to get there" (ARG, 2002, p. 2). This definition is similar to Sadler's (1989) idea about feedback and FA, where the author emphasizes that feedback, to be formative, should bridge the gap between the initial and the target level. Apart from being narrowly principled in feedback, AfL is seen as "any assessment for which the first priority in its design and practice is to serve the purpose of promoting students' learning" (Black et al., 2003, p. 10). The authors of this definition are ostensibly going in the direction of how AfL should be implemented instead of looking at the purpose for which it is used. In an eminent conference in New Zealand, which gathered eminent scholars around the world, the participants issued an updated definition of AfL, reported by Klenowki (2009), accentuating its dynamic nature as being integral to teaching and learning, saying that AfL is "part of everyday practice by students, teachers and peers that seeks, reflects upon and responds to information from dialogue, demonstration and observation in ways that enhance ongoing learning" (Klenowski, 2009, p. 264). This is to say that AfL should be an integral part to everyday learning and teaching and not just a separate impromptu activity to be included or merely added.

Focusing on increasing the possibility of learning with AfL instead of guaranteeing learning occurrence and fostering effectiveness at the expense of progress, Laveault and Allal (2016) define AfL as "the collection and interpretation of assessment information whose intentional use enables teachers and students, acting individually or interactively, to reach decisions that have a positive impact on instruction and learning" (p.7). The authors emphasize that AfL shares the interests of teachers and students without either being in a higher position, and this confers a kind of mutual interaction in order to collaborate and discuss the status quo emerging in the classrooms. On the other hand, Wiliam (2020) has recently defined AfL simply as any assessment used essentially to address learning improvement instead of measuring it. In fact, this does not appear an easy or trivial definition because if we look at the ultimate goal, improving learning and trying to achieve it, here we can see how difficult it is the mission. To improve learning, one could simply focus on improving teaching strategies, motivating students, activating interaction in classrooms to create a community of practice, looking after setting self-regulated learning, and fostering toward autonomy and lifelong learning.

In a few words, AfL is not constrained with any specific definition and that makes it debatable. However, we can say that AfL is the assessment that is used to help learners improve and enhance their learning, alerting them about their weaknesses and strengths to work on the former and exploit the latter to facilitate their progress, after they know where they are, where to go, and how to go there. In addition to that, AfL can help them spot and identify their progress along a continuum that bridges those stations.

1.3 Assessment for Learning and Self-regulated Learning

Self-regulated learning is not a concept that is tightly linked to assessment, but it is used in many other disciplines like sociology, psychology, computer science, and others. The starting point of self-regulated learning in the literature was remarkably marked by two widely read papers: the first one by Crooks (1988) in New Zealand and the second one by Sadler (1989) in Australia. Both authors discussed how to provide opportunities for learners to learn through classroom assessment. Indeed, it is only in the first decade of the twenty-first century that both self-regulated learning and AfL become well-established fields of educational research (Panadero, Andrade, & Brookhart, 2018). In education and psychology, self-regulated learning is linked to cognitive, social and motivational mechanisms informed by learning theories including reinforcement in behaviorist theory, equilibration in Piaget's constructivism, feedback loops in cognitive models, and social mediation in sociocultural and social constructivist approaches (Allal, 2010).

In fact, self-regulation is seen as an aspect of AfL which helps students be responsible of their own learning (Wiliam, 2010). The connection between FA and self-regulated learning has been largely communicated through the works of Black and Wiliam (1998) in which they emphasized enhancing self-regulated learning as one reason for implementing FA. In other words, self-regulated learning was emphasized as a main goal of FA and it is basically linked to self-assessment. It is generated through the practice of assessment and feedback issued for the purpose of monitoring and evaluating the quality and impact of their and others' work (Nicol, 2010). For the feedback function, there is a relationship generated between what is called internal and external feedbacks and this is seen pivotal for students' development of self-regulated learning (Winne & Butler, 1995). To further understand this point by Winne and Butler (1995), see Section 1.5.1 paragraph 3.

Simply put, following Allal (2010) and Zimmerman and Schunk (2011) regulation process includes setting goals, monitoring the progress toward the preset goals, exploiting feedback generated from monitoring, and achieving those goals. As a result, self-regulated learning, as it combines the learning strategies, helps students learn effectively, motivates them, and teaches them how to control themselves (Andrade & Heritage, 2018). When being able to self-regulate their learning, learners can move backward and onwards and can develop an expertise (Bransford, Brown, & Cocking, 2000), which helps them to promote inner abilities and critical thinking to generate further progress. This is why self-regulated is seen highly fostered by AfL. For further understanding, we direct the reader to Section 1.7.1 to understand how self-assessment can enhance self-regulation learning.

1.4 Assessment for Learning Implementations

As a consequence of the various definitions related to different strategies and aspects of AfL incarnations, the implementation challenges are also diverse and that would be due to the "oversimplification" of the concept of AfL (Laveault & Allal, 2016, p. 6). In this section, we first look at the implementation's geographical spread and discuss how easy or difficult AfL is to implement, focusing on how and why this results.

Despite being founded on a set of principles, informed by a set of strategies, and interwoven with various aspects, AfL is not widely dispersed geographically- we mean here its practical implementation in classrooms- even if Laveault and Allal (2016) consider it to be so. The authors focused on some parts of the world and disregarded others. They drew on works such as the one by Hopfenbeck, Petour, and Tolo (2015) which described large-scale implementations at both national and regional levels in Norway, emphasizing particularly both the effects of assessment for purpose of accountability and SA within AfL implementations. Another work by Birenbaum, DeLuca, Earl, Heritage, Klenowski, Looney, Smith, Timperley, Volante, and Wyatt-Smith (2015) reported major issues in AfL implementation policy and practice in countries like Canada, New Zealand, USA, and Norway. However, in the Arab world, African and most Asian countries, none of the studies were taken into account and no reality about AfL implementations in classroom had been reported. For this reason, we consider that AfL is not well spread all over the world because the sample chosen was not valid for generalizability. We have deliberately cast light on regional implementation to warn researchers about the AfL research field and to state plainly that this topic is still in its early stages and requires further contributions and substantial investments.

The major problem with AfL implementations lies in the ambiguity that shadows this notion, and the potential benefits and the corresponding expectations of AfL that educators and learners have not been able to reap. To implement AfL, the following quotation by Laveault and Allal (2016) warns about some obstacles that might hinder AfL' effectiveness.

> While different incarnations of assessment for learning may be possible depending on the interactions between teachers and students, and among students, some forms of assessment for learning may not flourish or even be possible depending on the level of control which teachers and students are able to exert on the processes of regulation in their educational environment. The implementation of different forms of assessment for leaning is influenced by educational policies and assessment frameworks which shape the context for teachers' professional development and their collaborative learning about assessment for learning. Different combinations of external factors, with a variety of emphases on the processes of regulation, may lead to more or less successful, adaptive occurrences of assessment for learning. (p. 7)

Three major obstacles could be inferred from this quotation: the interaction in the classroom, the impact of the educational policies and assessment frameworks, and various combinations of external factors. Another emerging idea from this quotation is that there are different forms of AfL. This unequivocally hints at the different AfL combinations that could take place without limiting it to any particular form. For example, AfL can happen through implementing self-assessment in classroom to improve learning, or via teacher and self-assessment for instance. This offers some flexibility in operationalizing AfL for any classroom, going down the path that is fruitful. However, each form should be scrutinized in light of the three obstacles highlighted above which may prevent AfL from achieving its goal.

In order to best discuss the implementation issue, we need to take a closer look at the major principles underlying this concept. As discussed in Sections 1.1 and 1.2 AfL does not have a theory or theories which frame(s) it but instead is open to all contributions that might uphold its occurrence. The main contribution, which brings together a set of principles to provide an overview of AfL implementation, comes from the ARG (well-known the ten principles of ARG). Before discussing these principles, in what follows, we would like to point out that that these principles have existed for decades even though they are not clearly stated under the term AfL. The principles were established purposively to exploit assessment experiences to support student learning. For example, Rowntree (1987) suggested seventeen proposals for improving assessment practice in higher education, and Chickering and Gamson (1987) proposed seven principles for good practice in undergraduate education, referring to wider aspects of higher education. Following Draper (2007) and Nicol (2008), these principles could also be applied to assessment. We have mentioned these works to help the reader get a complete idea, but discussing them here does not contribute to the main aim of this research.

In fact, the principles that make AfL come into existence and spring up in the literature were carried out through Black and Wiliam's (1998) seminal work. These principles were published by the ARG, and that helped them to widely spread to a wider audience. The ten principles suggested by ARG are meant to guide classroom practices. In this regard, AfL should:

1. be part of effective planning of teaching and learning
2. focus on how students learn
3. be recognised as central to classroom practice
4. be regarded as a key professional skill for teachers
5. be sensitive and constructive because any assessment has an emotional impact
6. take account of the importance of learner motivation
7. promote commitment to learning goals and a shared understanding of the criteria by which they are assessed
8. provide constructive guidance to learners about how to improve

9. help in developing learners' capacity for summative assessment so that they can become reflective and self-managing
10. recognise the full range of achievements of all learners (ARG, 2002, p. 2).

The ten principles have benefited from a wide range of comments of various associations and researchers to be a further and advanced step toward modifying assessment practice to ensure the essential and necessary quality of learning experiences required for fulfilling the goals of education (ARG, 2002). They came to support the eminent works presented by *Black and Wiliam: Inside the Black Box*, and the follow-up work, *AfL: beyond the black box*. We attempted to distill these ten principles into five major ideas: (a) integrating AfL at core of the classroom, (b) involving students as active participants for further contribution and progress, (c) exploiting feedback, (d) reflecting on teaching practices, and (e) creating a community of practice for dual benefits: for the learner and for the group.

Indeed, we can observe that these ten principles provide a significant contribution to applying AfL for improved learning despite the presence of other attempts flowing in this context. We concisely present five attempts, which have also been undertaken in this context of AfL implementation. The first attempt was made by Gibbs (2006). He reviewed the conditions that make assessment works for learning and for this purpose he suggested that assessment should be students' main interest both in time and effort. He also proposed issues related to feedback effectiveness pointing to feedback frequency, timeliness, its relation to the assessment criteria, and the influence of feedback on future learning. The second attempt proposed by Wiliam and Thompson (2007) emphasized five key aspects in AfL which are (a) clarifying and sharing learning intentions and success criteria, (b) focusing on questioning and classroom discussions for further understanding, (c) providing feedback that feedforward student learning, (d) activating students to build their own learning, and (e) encouraging students to be interactive one to another. The third attempt by Boud and Associates (2010), in light of the assessment reform, suggests that (1) assessment should engage students in productive learning, (2) feedback has to be used for improvement sake, (3) teachers and students should interact and be partners in learning and assessment, and (4) assessment should be positioned at the center of course design. Another contribution made by Sambell, McDowell, and Montgomery (2013), the fourth attempt, where the authors suggested four features of AfL. They pointed to (a) balancing appropriately between summative and FA, (b) opting for authentic complex assessment task, (c) self-evaluation activities enriched with both formal and informal feedback, and (d) offering opportunities in practice for building confidence.

Recently, based on the works proposed about implementing and integrating AfL Careless (2017), the fifth attempt, tried to synthesize the main AfL aspects for further implementation accompanied with their operationalization in Table 1.2.

Table1.2 *Synthesis of Main AfL Strategies and Processes (Careless, 2017, p. 7)*

AfL strategies	**Illustrative implementation processes**
Productive assessment task design	Alignment with intended learning outcomes
	Authentic assessment
	Integrated and coherent assessment
	Collaborative writing through wikis
Effective feedback processes	Integrated guidance and feedback
	Students generating and seeking feedback
	Closing feedback loops
	Technology-enabled feedback dialogues
Developing student understanding of the nature of quality	Students generating and/or decoding criteria
	Applying criteria
	Analysing and discussing exemplars
	Online dialogue about exemplars
Students practising making judgements	Providing peer feedback
	Receiving peer feedback
	Self-monitoring work in progress
	Online facilitation of peer interaction

Based on the works presented above, the ten principles of ARG (2002), Gibbs (2006), Wiliam and Thompson (2007), Boud and his colleagues (2010), Sambell, McDowell and Montgomery (2013), and Careless (2017), we have tried to highlight the major similarities and differences between them to arrive at a reasonable synthesis. The common aspects shared with all those works are the need for feedback and putting assessment at the center of the classroom to be able to call AfL as such, whereas the differences we spotted could be summarized into three major points. The first difference is how to activate feedback. For example in the work by Gibbs (2006) the author focused on frequency, timeliness, effectiveness and activating feedback for future learning while in the work by Wiliam and Thompson (2007) the authors pointed only to feedforward. In the work by Boud and colleagues (2010), feedback is for improvement without specifying how that should be, and for Sambell, McDowell, and Montgomery (2013) feedback has emerged in self-evaluation activities where it should be formal and informal. The second difference is the responsibility related to AfL. Some major questions could be asked like whose is the responsibility? Are both teachers and students concerned with it? Is this responsibility shared equally or not? The last difference is the focus or not on the goals, either short, long, or both, and their differences. For example, The ARG's (2002) ten principles focused on goals while the other works did not owe much importance to that.

To conclude, implementing AfL is a difficult undertaking since various requirements and factors must be taken into account. The scarcity of the literature on AfL in higher education makes it impossible for scholars to make definitive statements about AfL strategies (Careless, 2017), but effective implementation is based primordially on making assessment an integral part of the classroom and creating space for critical interactivity between the teacher and students as well as among students. Furthermore, AfL should provide students with opportunities to react on their own in order to profile themselves and monitor their progress. Most importantly, AfL should be enjoyable and motivating in order to promote much self-regulation, self-confidence, autonomy, and lifelong learning. Finally, AfL can achieve all that since it is basically fostering focus on offering detailed and constructive feedback.

Part II. Classroom-based Assessment: Theoretical and Conceptual Framework

In this part, we shall define classroom-based assessment, focusing on the components that contribute to shaping it as an entity which offers students a space to interact effectively. Toward this aim, feedback as a vector that can nail students' needs and promotion, rubrics as assessment tools, self-, peer and TA as assessment practices that can promote classrooms, and portfolios as tools that document students' progress will be discussed accordingly.

1.5 Feedback

Feedback emergence in education can be linked to the behaviorist paradigm which lies in the process of stimulus-response chain. However, recently it has been given due attention in the literature about education because it has been acknowledged by the majority of scholars to be pivotal, undeniable, and having a crucial role in classrooms. It is classified with classroom discussion and teacher clarity with the top ten rankings of 150 factors of over 50, 000 research studies (Hattie, 2012).

In this thesis, we have addressed the significance of this concept, how and when it is effective, what roles teacher feedback play, how significant anonymity is in PA, and what relationship feedback has with AfL.

1.5.1 Feedback Defined

Feedback has been defined variously; meanings range from simply stimulating (the behaviorist approach) passing by correcting to offering information for further improvement. Dating back more than a century ago, Thorndike (1913) sees that feedback can be positive (reinforce) and negative (punish). However, in the 1970s Kulhavy (1977) asserted that the behaviorist view was a hasty decision because the laboratory conditions are not the same as the real in-classroom ones, and feedback may be accepted, rejected, or even modified. For this purpose, Kulhavy (1977) suggested that if feedback is a generic term and a unitary variable, its form or composition would range from simple yes-no answer to "substantial corrective or remedial information"; along the continuum, the feedback function varies from informing students about correctness to generating new instruction (p. 212). Later, three broad meanings of feedback have been examined by Kulhavy and Wagner (1993), which are motivational, reinforcement, and/or informational. Feedback is motivational when it increases the general behavior such as revising and writing. Feedback has reinforcement meaning when it can punish or reward, and it is informational when it is applied by students to change a performance in a specific direction. With reference to writing, Nelson and Schunn (2008) see that these three meanings, motivational, reinforcement, and/or informational, would all be important. However these three functions may not be carried out by one feedback information and may not be appropriately understood by all students. For Hyland (2003), feedback refers to any comment, underlining, or correction, written by teachers and made on student's text. Even when offered on writings, feedback cannot be merely written, and even if it is, this does not ensure its usefulness.

In their notable work, Hattie and Timperley (2007) conceive feedback as follows in an attempt to establish a more or less flexible definition of feedback.

> Feedback is the information provided by an agent (e.g., teacher, peer, book, parent, self, experience) regarding aspects of one's performance or understanding. A teacher or parent can provide corrective information, a peer can provide an alternative strategy, a book can provide information to clarify ideas, a parent can provide encouragement, and a learner can look up the answer to evaluate the correctness of a response. Feedback thus is a "consequence" of performance. (p. 81)

This definition seems lacking detail and rigor, making it shallow and uninsightful. For example, a book is not feedback. It is just a book. It only becomes feedback when the reader uses the book for a particular purpose. Another point to emphasize is that teachers provide corrections while students provide alternative strategies. This would pose some questions like: Are students unable to correct? Which is better: correcting or deploying a strategy? What does a strategy include? Is teacher feedback limited to just corrections? And so forth. A number of scholars see that feedback definition appears narrow as they focus on certain features such as establishing its route of delivery, function, or even type. For instance, Kepner (1991) defined feedback as an input that informs students if a particular instructional answer is correct or incorrect for the purpose of revision and correction. We are opposed to confining feedback to such definitions because it would deny feedback the right place to occupy because feedback is currently used to inform and direct cognitive processes for larger comprehension and further knowledge building.

From self-regulation perspective, Butler and Winne (1995) argue that feedback is identified as an “inherent catalyst” because when students monitor their own engagement with tasks, they generate internal feedback to engage with the external feedback provided by the teacher (p. 246). Both internal and external feedback sources provide referent points for ongoing refinement of goals and plans as well as for processes destined for undertaking the task. In writing, students use external feedback which is provided in the form of comments, corrections, guidance, advice and even scores to reflect on their improvement regarding predefined criteria. This could reduce the misunderstanding of internal perceptions related to writing to meet realistic expectations. Nicol and Macfarlane-Dick (2006), on the other hand, defined feedback to fulfill the function that “anything that might strengthen the students’ capacity to self-regulate their own performance” (p.206), in hope of accelerating their learning and meeting their goals.

A very important conceptualization of feedback which is still efficient is the one suggested by Ramaprasad (1983) and Sadler (1989). They see that feedback is linked to reducing the gap between the current and the target level. Rather than being defined in terms of its content, Ramaprasad (1983) prefers defining it in terms of its effect as follows: “feedback is information about the gap between the actual level and the reference level of a system parameter which is used to alter the gap in some way” (p. 4). For Ramaprasad (1983), feedback is only called as such when it succeeds in altering that gap. To that end, Hattie and Timperley (2007) proposed some ways for closing that gap such as affective processes including increased effort, motivation, or engagement. They also see that

> the gap may be reduced through a number of different cognitive processes, including restructuring understandings, confirming to students that they are correct or incorrect, indicating that more information is available or needed, pointing to directions students could pursue, and/or indicating alternative strategies to understand particular information. (Hattie and Timperley, 2007, p. 81)

This can clearly suggest that reducing the gap is not only linked to the materials used to measure the responsiveness degree to feedback, but also does internal processing play a major role in the way of perceiving that feedback. On the other hand, both teachers and students might be found in a competitive position to be either a principle actor in maximizing the degree of feedback effectiveness. To achieve that, feedback in this context is provided for two target audiences, teachers and students; the former exploits it in respect to diagnosis, readiness and remediation to make planned decisions while the latter exploit feedback to manage the strengths and weaknesses of their performance to either recognize, modify, reinforce, improve, or exploit them (Sadler, 1989). In other words, feedback is information provided to learners to use as a medium "to confirm, add to, overwrite, tune, or restructure information in memory, whether that information is domain knowledge, meta-cognitive knowledge, beliefs about self and tasks, or cognitive tactics and strategies" (Winne & Butler, 1994, p. 5740).

In a nutshell, feedback definitions are so diverse and broad that they do not share common components because of the various purposes, ways, types, and goals for which feedback is used. Feedback has started to be viewed simply as saying yes or no and correcting mistakes, but recently it has been destined for the sake of targeting cognitive processings. Bringing all visions closer to the purpose of this study, the feedback emphasized is the one that is directed to help students improve their writing within AfL. This is to say that feedback is linked to effectiveness regardless of its form or type. Therefore, feedback can be defined as the information that helps students progress; it not only corrects but also informs, describes, prescribes, instructs, and assesses.

1.5.2 Teacher Feedback

Despite the importance of teacher feedback in the writing process as being a way to help students refine their writings more than other types (Lee & Schellart, 2008), many studies about this type of feedback, i.e. teacher feedback, have end up at contradictory findings concerning students' preferences to feedback. For example, Hyland and Hyland (2006) found that students prefer their teachers' feedback while Jacobs, Kahn, Stralka, and Phan (1998) found that students prefer instead peer feedback. In this study, what matters is not feedback preference, i.e. which type, but rather feedback quality content and appropriateness. Feedback quality remains important in improving students' writing, and teachers are in a "difficult situation with respect to providing student feedback" (Orsmond, Merry, & Reiling, 2002). In relation to this reality, many factors could be involved in providing effective teacher feedback. The aspects under focus are the relationship between teacher and students, involving students in feedback process, and targeting appropriate feedback type and form.

- **Relationship between Teacher and Students**. This relation is playing a key role in how teachers give feedback to students and how they respond to this feedback (Blakeslee, 2001). To be effective, both parts, students and teachers, should have a common understanding of how feedback should be implemented and exploited. Teachers have to ensure that the feedback they give cannot be easily skipped or dismissed by students to make them revise their writing in a "consistently, narrow, and predictable way" (Orsmond et al., 2002, p. 233). However, students' perceptions of teacher feedback, including teacher credibility, can be impacted undesirably by teacher response to their written work (Lee & Schallert, 2008; Witt & Kerssen-Griep, 2011) which may lead feedback given to students to threaten their freedom to act or to interpret it negatively (Witt & Kerssen-Griep, 2011). It appears that the credibility of the source giving feedback should not only depend on providing accurate feedback but also has to be careful when, how, and why providing that feedback.
- **Involving Students in Feedback Process.** Involving students may be another successful strategy to help both, teacher and students, gain insightful understanding of the mutual interaction between teaching and learning. In this regard, teachers should try to deliver effective feedback and students should try to respond to that feedback accurately, appropriately, and fruitfully. To clarify how feedback works in writing, Goldstein (2005) points to that writing revision is a process with multiple and interactive factors "mediating each other, through a cyclical process within which these multiple student texts and teacher commentary texts are created" (p. 24). However, including only written commentary texts and multiple drafts is not enough to call a process revision. Hence, including oral feedback and guidance can promote revision and put it on the right track. All in all, involving students in a revision process can facilitate

their “long-term improvement and cognitive change” (Reid, 1993, p. 229), help them increase their revision skills (Ferris & Hedgcock, 2014), enhance in them motivation (Sommers, 2013), establish and maintain connections between writing instruction and students’ academic profile (Beaufort, 2007), and promote self-regulated writers (Andrade & Evans, 2013).

- **Targeting Appropriate Feedback Type and Form.** Providing effective feedback is not that easy. To uncover this complexity, a lot of experimental studies had been conducted discussing and comparing the effects of different types of feedback on writing performance (e.g. Bitchener, 2008; Ferris, 2006) but without final conclusive results. For example, Hyland (2003) found that most of teacher feedback was based on language accuracy, and early research in L2 writing started to be critical of that kind of feedback as focusing on grammar is not helpful and discouraging in the same time. Feedback focus on language accuracy has started to be effective, then less effective (Lalande, 1982; Kepner, 1991) and after ineffective in most cases (Truscott, 1996). As a response to that, Baleghizadeh and Dadashi (2011) found that indirect teacher feedback is more effective than the direct type in improvising actively classroom learning.

Clearly put, there is no specific feedback type which has been proved to be effective. However, teacher feedback, in its form and content, may impact students’ writing progress and their personal academic traits like self-esteem, confidence, and motivation (Hyland & Hyland, 2006; Witt & Kerssen-Griep, 2011). For example, teacher written comments may connect teacher to students at an individual level (Ferris, Pezone, Tade, & Tinti, 1997), and may affect students’ writing and attitude. In light of that many questions can be raised like: Is the feedback provided related to the context? What kind of relations does teacher feedback promote? To what extent is teacher feedback provided effective? In this regard, Reid (1994) sees that written comments are usually seen as written texts without examining the context or the relationship developed between teachers and students, and this ,in turn, might impact feedback interpretation and production and its overall contribution (Reid, 1994).

Another point to emphasize is the negativity of the feedback. It is found that teacher negative feedback may negatively impact learner identity, self-esteem (Carnicelli, 1980), and motivation (Värlander, 2008), and that may damage the learning process (Shvidko, 2020). In fact, we see that excessive negative feedback would result in the aforementioned disadvantages, but sometimes this kind of feedback may assist in informing learners about their mistakes if they are committed repeatedly, continuously, and carelessly. Despite being positive or negative, Hyland and Hyland (2006), their turn, see that teacher feedback could either block students' voice and impose teacher requirements or empower students to produce writings that meet the standards in a specific setting. For this reason, Shvidko (2020) argues that delivering feedback for teachers is a challenging task as, to some extent, teachers should not only target students' negative aspects in writing but also the positive ones. As a result, balancing the dual types of feedback, negative and positive, could lead to much more better learning instead of firmly focusing on either type alone.

To conclude the discussion about teacher feedback, we can say that teacher feedback should be present but not occupy all the space given to classroom feedback. In AfL, teachers should be guides and helpers and their feedback should guide and help. When necessary, the feedback may include corrections, advice, suggestions, warnings, instructions, prescriptions, and encouragements. It is also preferable to be of mixed types rather than being limited to one specific kind in order to target different preferences, needs, and abilities in the classroom.

1.5.3 Anonymity in Peer Feedback and Peer Assessment

Before launching the discussion on the theme, it seems of paramount importance to shed light on two important points. The first one is to draw the reader's attention to the fact that we did not find any study discussing anonymity in PA. All the works found linked anonymity to peer feedback, which they interchanged with other concepts like PA, peer review, peer comments and peer evaluation. The second point is that PA is not put synonymous to peer feedback in this thesis despite being treated as so in the literature. To review anonymity, we first defined it, focusing on its positive and negative impacts, and then we discussed its implications.

Despite the potential benefits reported about involving students in providing feedback in many studies in the literature, this did not remain without any problem. In this context, many researchers suggested that students can easily be biased and cannot be honest because of gender difference, race, friendship, personal preferences and interpersonal relationships (Zhao, 1998; MacLeod, 1999; Nilson, 2003). To remediate this problem, many scholars suggested opting for anonymity. This concept is defined by Marx (1999) as the condition in which the sender of the messages is not known either in terms of their personal identity or their relevance to a social context. The importance of anonymity lies in creating a psychological state of deindividuation where the person gets involved in the group with less individual identity (Jessup, Connolly, & Tansik, 1990) to exercise in the community freely and spontaneously. Deindividuation, in its turn, is defined to be the existence of individuals in groups without thinking of the other members as individuals discarding any feeling of being singled out (ibid). As a result, it is seen that in anonymous situations people are more honest and less anxious while expressing their opinions which could drive them to be more critical (Zhao, 1998; MacLeod, 1999).

Indeed, anonymity is seen practiced and approached differently. It can be directional, one part only-either assessor or assessee- is known, or bidirectional, both parts are unknown (Panadero & Alqassab, 2019). It can also be provided in various ways such as using pseudonym or random assignment without identification (Yu & Wu, 2011) as it can be linked to the purpose of PA (Panadero & Alqassab, 2019). Depending on the goal, PA may be used for summative or formative purposes, and anonymity may be extremely valuable in this context, of PA for AfL purposes, especially if it is not misunderstood and misused. Also, it might not be useful even if it is used correctly understood and used. This is to say that it is a debatable issue. Strijbos, Ochoa, Sluijsmans, Segers, and Tillema (2009), for example, asserted that the value of anonymous PA is observed in summative or high-stake implications, although this is not clearly articulated. The importance of anonymity can lie in providing fair assessment by preventing the impact of social effects, scaffolding the initial steps of PA to make students more confident (Rotsaert, Panadero, & Schellens, 2018), and making PA process more powerful, fairer, and safer (van Gennip et al., 2009; Panadero, 2016). In the same context, Panadero and Alqassab (2019) proposed summarizing anonymity effects into two major outcomes: (a) cognitive outcomes, namely the content of peer feedback and PA accuracy and achievement or performance, and (b) social-affective outcomes, including students' perceptions of PA, interactivity in the group such as frequency of interactions, and social conflicts such as peer pressure and psychological safety.

From an empirical evidence angle, the importance of anonymity in PA is reported as being various and conflicting. Starting with the positive outcomes, comparing feedback provided anonymously to the one identifiably Lin (2017) concluded that the former group, the groups which provided feedback anonymously, contributed with more cognitive feedback such as suggestion and extension and learned a lot from PA in addition to perceiving peer comment to be less fair, than did the former group. While for the identified peer feedback, students in this group offered more affective feedback, namely supporting, and more metacognitive one, namely reflective comments. On the other hand, Lu and Bol (2007) and Lin (2017) found that APA helped students to outscore the group who received identified peer feedback whereas Yu (2012) found no difference between the two groups. For Zhao (1998), online APA was found providing more critical feedback, and for Guardado and Shi (2007) online APA was found offering more honesty and directness in the feedback in addition to being more critical. In a study conducted by Rotsaert et al. (2018) investigating the impact of using APA gradually on the quality of peer feedback, the authors found that peer feedback quality had increased significantly from session one to session two but thereafter it remained stable.

Still under concern of anonymity effects in PA, Wachwa, Schulz, and Mann (2006) found that students in online APA had provided significant feedback, either positive or negative, which tended to be cognitive-oriented while in the identified PA students provided less-deep cognitive comments. Finally, in their review of fourteen studies, all empirical, using PA interchangeably with peer feedback, peer review, peer evaluation and peer grading based on a control or within group design, Panadero and Alqassab (2019) found that APA helped student to value PA, providing more critical feedback, increasing self-perceived social effects with slight more performance in higher education without using a lot of aids. In addition to that, the authors made several recommendations, such as taking the instructional context and learning goals into account when introducing anonymity in PA, as well as pointing out the lack of empirical evidence and encouraging further study using more complicated research designs.

Despite the positive effects generated by anonymity in PA, this concept is not without contradictory views. In line with this idea, the findings about APA have taken various directions (Panadero & Alqassab, 2019) and have not always been found positive (Yu, 2012). According to some researchers, anonymous feedback is more negative (Zhao, 1998; Lu & Bol, 2007), less cognitively oriented, worse in quality, less valuable (Zhao, 1998), and less fair than the one delivered identifiably (Kaufman & Schunn, 2011). In peer grading, recent meta-analysis found that anonymous peer grading is less reliable than the identified one (Li, Xiong, Zang, Kornhaber, Lyu, Chung, & Suen, 2015), and Panadero and Alqassab (2019) suggested that the presence of peer grading in PA might moderate the effect of anonymity on the accuracy of PA. (Panadero & Alqassab, 2019). On the other side, the same authors shed light on some dilemmas that could appear when introducing anonymity, for instance focusing on peer comments/grades reliability, using grades to support comments, and training and building a community to support the culture of feedback. From an anonymity implementation perspective, some major problems have also been spotted. The major problem reported about anonymity is loafing both physically and cognitively (Jessup et al., 1990). This might be due to approaching anonymity in a simplistic way, disregarding the complexity of social interactions and the psychological side of both the assessor and assessee, which can generate various and unpredictable behaviors toward anonymity.

To summarize, anonymity should be opted for only when it is seen effective because it might not always yield positive outcomes and it should not be seen as the only solution when problems appear in PA. Therefore, it could be dependent on the decisions taken by the teacher or researcher regarding the aspects included in PA (Panadero & Alqassab, 2019). In this study, we are concerned with bidirectional anonymity as this latter is seen adequate to the current study. This could help adjust students' learning identities and discard the sense of subjectivity that students may suffer from. Nevertheless it is used for assessment learning purposes and there would be a need for social interactivity, we see that anonymity could be less useful with advanced levels than beginners and low intermediate levels. This in any case could in fact undermine students' abilities and capacities to interact and feed each other writing. However, being newly introduced to a context where AfL takes place for the first time, anonymity could be taken as a prerequisite to avoiding social conflicts and unpredictable reactions like the impact of friendship, gender difference and differences in capacities.

1.5.4 Feedback and Assessment for Learning

In light of the discussion in sections 1.5.1 and 1.5.2, despite inconsistent results on the effectiveness of various types of feedback supplied in various methods, feedback, in particular, plays a key role in learning and teaching in general, and writing in particular. Whatever the type and form it takes (written and/or oral form, direct and/or indirect), whoever delivers it (teacher and/or peers), and for which purpose is used (corrective, descriptive, prescriptive, etc), nearly all scholars agreed on the need of providing feedback for effective learning. In this part, we will look more closely at the significance of feedback in AfL.

Feedback emerges to take a central position in helping students progress in their learning in higher education as assessment is partly involved in shaping learning (Sadler, 2010); hence, summative feedback, which is used to grade the final products in writing, has been recently replaced by formative feedback (Kathpalia & Heah, 2017). Formative feedback in process-based writing classes seems favored as it provides guidance in developing composing skills in addition to shaping future writing

(Hyland & Hyland, 2006). This is to say that feedback is not for merely correcting but making students acquire skills for further critical writing whatever the used teaching approach.

The relation of feedback to AfL is seen in shifting from one-way feedback provision to multi-ways delivery. The former implies unidirectional feedback provision from teacher to students while the latter suggests that feedback is provided from teacher to students, from students to teachers, and from students to students. In this context, feedback practices have in fact shifted lately from teacher written feedback to peer feedback and oral feedback in classroom teacher-student interactions, namely conferencing in addition to peer interactions to displace the emphasis from mechanical accuracy to developing meaning in the text (Kathpalia & Heah, 2017) or properly the context. Clearly, feedback is a key aspect to make AfL successful (Boyd, Green, Hopfenbeck, & Stobart, 2019) but definitely not all of it (Wiliam, 2018). This is true only if together FA and feedback empower students to become self-regulated learners (Careless, 2007).

The importance of involving students to be feedback providers is very crucial. In line with this, many authors found that students are more effective when invited to provide feedback (e.g. Bangert-Downs, Kulik, Kulik & Morgan, 1991). For example, Green (2019) sees that student feedback to teachers is seen a part of the most useful feedback provided in classroom, and only the best teachers react to that feedback to improve the quality of their teaching. This, to some extent, might not be accepted by teachers who might resist that as they consider themselves more knowledgeable and more expert than students, but these ideas should vanish with the emergence of AfL. In the same vein, Hattie (2009) concluded that

> feedback was most powerful when it is from the student to the teacher…When teachers seek, or are at least open to, feedback from students as to what students know, what they understand where they make errors, when they have misconceptions, when they are not engaged-then teaching and learning can be synchronized and powerful. Feedback to teachers makes learning visible. (p. 173)

This can denote that students should be involved in the feedback provision process as this could effectively boost their critical thinking toward a realistic understanding of their mission in the learning atmosphere; to learn when, why, and how to engage and for what purposes. In agreement with this, Wiliam's (2011) summary of the research on feedback concluded that "feedback should cause thinking" (p. 127). This means that students should be able to understand how feedback works best in order to understand why they are there in the classroom. While understanding, acting, and managing the feedback received, learners develop "a more questioning and reflective attitude to learning, increasing motivation and resilience, and ensuring deeper learning" (Boyd, Green, Hopfenbeck, & Stobart, 2019, p. 4). Given such a useful contribution of feedback processing, feedback should help students understand their learning goal, the ways to achieve it, and how to bridge the gap between the initial and the target status (Sadler, 1989). Moreover, Brookhart (2005) sees that common classroom assessment aims include providing students with feedback for their studying, making instructional judgments, issuing marks, and counseling students about extra courses.

In a nutshell, a typical classroom enmeshed by AfL purposes should provide students with feedback, regardless its type, form, and content, in a way to be purposive and effective to comply with students as well as institutions' goals and objectives.

1.6 Writing Rubrics

As the literature on rubrics, in particular for grading purposes, is substantial, we would like to explain how we approached reviewing this concept in relation to the present study. First, we did not rely on one database to gather the needed documentation. Second, the review is not limited to writing rubrics, but various studies on using rubrics in other fields have been also consulted to enrich the discussion. Third, we did not depend just on research conducted in higher education contexts; pre-university level studies are also included because they show substantial contributions, some of which were deemed remarkable in the field of rubrics. Reviewing is not limited to rubrics in EFL context because this field, at the time we carried this study, was not well unearthed and many studies were needed to bring practical contributions.

The basic reality is that rubrics are addressed for assessing both process and product performances (Brookhart, 2013). Even though they have settled in the US assessment culture for decades, recently many scholars such as Reddy and Andrade (2010), McConnell (2013), and others have encouraged the importance of expanding their use outside the US to see whether they match or mismatch what had been concluded at in the USA region and other parts of the world. To review this concept, we shall discuss the conceptualization of rubrics as assessment tools eliciting their use and effectiveness, or their uselessness and emptiness, in guiding students to improve, or not, their English writing performance. In addition to that, we shall clarify some misconceptions about validity and reliability issues.

1.6.1 What Is Rubric?

The basic meaning of this concept, according to Oxford Dictionary, was derived, in the mid-fifteenth century, from the Latin word 'ruber' which means red. It first meant the major headings of a book following Christina Monks who reproduced a sacred literature starting each heading with a large red word. Later, the word rubric gained another meaning among educators which meant describing the rules of their scoring. As it brought much importance, Andrade (1997) asserted that the term rubric challenged the dictionary definition and established a remarkable existence by itself, and for this reason she continued using it.

To define rubrics, various definitions are examined in order to reach a synthesis of a clearer idea to conceptualize what rubrics truly mean in accordance with this study's goal. In the online Merriam-Webster dictionary, rubrics are defined to be "a guide listing specific criteria for grading or scoring academic papers, projects, or tests." (p. 3). For Brookhart (2005; 2013), rubrics are "scales, usually short ones, constructed to rate the quality of student work along a series of performance levels described under a criterion." (2005, p. 8), or *"coherent set of criteria for students' work that includes descriptions of levels of performance quality on the criteria"* (2013, p. 4. *Author's italics*). In writing, Andrade (1997) argued that rubrics are scoring tools that "lists the criteria for a piece of work, or "what counts" (for example, purpose, organization, details, voice, and mechanics are often what count in a piece of writing); it also articulates gradations of quality for each criterion, from excellent to poor” (p. 1).

In light of these definitions, rubrics have first appeared to be used for evaluation purposes, but more recently they are no longer used exclusively for that purpose. The dictionary definition does not appear to be fully detailed to meet the meaning a rubric has acquired recently as many questions could be asked such as in what order should the criteria be listed? What does it mean to list; does it mean just making a list or having principled procedures for listing? What fundamentals should they focus on to select this list? Why is it worth including a given criterion while not others? Is the rubric just for grading? So, the dictionary definition is, therefore, too general because checklists and prompts, for example, can also be included.

The distinction between the two definitions offered by Brookhart (2005; 2013) is that the first was intended for pure scoring end, whilst the second was discarded by the author since it was not deemed required or primordial. The author of the earlier definition noted that the rubric is short, but the latter is not. This implies that when it comes to creating rubrics, length is not an issue. Within the same perspective, some authors suggest that rubrics are limited to a certain number of pages, while others ignore this dimension. For instance, Andrade (1997) and Popham (1997) estimated rubrics to be one or two-page document. However, O’Donnell, Oakley, Haney, O’Neill, and Taylor (2011) posit that a useful rubric helps “on focusing the task without shuffling through pages of papers and inundating the faculty members (non-teachers like doctors, dentists and so on) (O’Donnell et al., 2011, p. 1166). This is suggested in the dentistry field, but the reality could also be the same in education as bothering students with lengthy rubrics might disorient or prevent them from progressing especially when student are not at an advanced level. But if that is seen necessary, long rubrics can be developed without any problem. For example, the Education Northwest (2018) 6+1 Trait® Writing Rubric is twelve (12) long pages large, yet it is highly recommended to be used. Its significance stems from its detailed content.

Having discussed the meaning and definitions of rubrics, we will now take a closer look at their main components.

Ideally, Popham (1997) sees that the three main aspects a rubric should include are “evaluative criteria, quality definitions, and a scoring strategy.” (p.72). The three aspects should co-exist in the rubric in a way that makes it as a whole effective. In this context, Cox, Morrison, and Brathwaite (2015) posit that it is of paramount importance to give “clear concise performance criteria and provide a forum in which students can create their own learning opportunities” (p. 27). What needs to be born in mind is that some criteria, such as voice, are not in fact recommended because they are ambiguous and misleading. This latter has been found omitted from many rubrics because it is thought too difficult to define. In agreement with this, Murray (2004) suggests that “voice is the quality, more than any other, that allows us to recognize excellent writing” (p. 21). As a result, it has been difficult to teach and establish as a unique criteria, as it appears to be gained via experience and a deeper comprehension of the cognitive perceptions of writing composing processes.

As for description levels, called also descriptors and performance levels, they have to be “specific, observable, and measurable …define expectations at each level of performance for each criterion” (Office of Institutional Research and Assessment, 2017, p. 3). They have to communicate the quality levels according to the user’s understanding level and to meet the expectations that have been already pre-set. For performance levels, their number is not fixed yet; it may range from four to six (Arter & Chappuis, 2007), or three to five (O’Donnell et al., 2011). We see that overall three to six levels are permissible, and less than three may dilute the meaning rubrics and may cause it to lose their significance. More than six levels may overload the rubrics, making them confusing and annoying, and it may also lose its usefulness even if they are properly developed. What counts most is that the descriptors should be appropriate and understood to both teachers and learners (O’Donnell et al., 2011). Concerning the scoring strategy, we found that it is not required, especially when using rubrics for learning or qualitative assessment. On the other hand, this depends on the scoring objective itself; whether it is designed to contribute to better learning or to account just for the progress made, i.e. for pure summative scoring goals. This is why some researchers did not emphasize including this feature in rubrics to demonstrate that rubrics may be used for purposes other than scoring.

Basically, it is important to bear in mind that one type of rubric is not inherently better than another, but a specific format can best work for specific purposes (Montgomery, 2000). For that reason, it is up to educators, teachers, or researchers, in agreement with students, to decide what to include in their rubrics: the criteria, the descriptors and the scoring strategy or the criteria and the descriptors without the scoring strategy. It is also their joint responsibility to define the structure of the rubric and the purposes for which it is used: for SA, or motivation and retrieval practices i.e. for AfL. As an illustrative example to clarify what has been previously discussed, in a study by Wang (2016) exploring the factors affecting the effectiveness of using writing rubrics, the author found that coverage and structure of the rubric, the descriptor performance quality, and score range were the main problems. The author used Jacobs, Zingraf, Wormuth, Hartfield, and Hughey's (1981) rubric. The participants of the study pointed to three main pitfalls regarding this rubric: (1) the narrowness of the rubric; they complained against including just five aspects which was not enough and therefore suggested other criteria to add such as students' style and voice of writing, (2) the quality definitions of the qualitative criteria, pointing to the subjective judgment featured in the descriptors, and (3) the analytic structure of the rubric. This example demonstrates that there is no standard or optimum rubric to utilize because each classroom has its own set of problems, and thus the rubric (s) that may assist to resolving those issues must be unique and appropriate to highlight and intervene in changing the situation under concern.

Overall, we can state that rubrics, as assessment tools, can grade, characterize writing performance, or combine all of these to serve learning goals. In addition to a scoring aspect, rubrics can be viewed as a set of criteria accompanied with degrees of performance stated accordingly. Rubrics are meant to be simple to use and understand, as well as to be purposeful, realistic, and effective. In this study, we are concerned with learning rubrics, which include the criteria matching descriptions and performance levels as well as a scoring procedure. These three components should work toward helping students progress in their writing. As for the rationale of structuring and designing the rubrics used in this study, they are fully discussed in Section 2.4.3.1 in research design chapter. The rubrics used in this study are intended to describe both the writing works done by the participants and their writing abilities. In this regard, Panadero, Jonsson, and Strijbos (2016) see that rubrics stress essentially the way of evaluating the product, the process, or both rather than offering a set of instructions about how to solve the task.

1.6.2 Rubrics Classified

Rubrics can be classified in terms of describing the performance or the criteria content. In terms of describing performance, rubrics can be either general or task-specific, while in terms of criteria content they are either holistic or analytic (Brookhart, 2013). These two classifications are basically the most prominent ones in the literature.

1.6.2.1 General vs. Task Specific. This classification received less attention than the other, holistic vs. analytic. The main distinctions between the two, general and task-specific, are displayed in Table 1.3, by Brookhart and Nitko (2008).

Table 1.3 *General vs. Task-Specific Rubrics. From Brookhart (2013, p. 8)*

Description of Performance: General or Task-Specific			
Type of Rubric	**Definition**	**Advantages**	**Disadvantages**
General	-Description of work gives characteristics that apply to a whole family of tasks (e.g., writing, problem solving)	-Can share with students, explicitly linking assessment and instruction. -Reuse same rubrics with several tasks or assignments. -Support learning by helping students see “good work” as bigger than one task. -Supports student self-evaluation -Students can help construct general	-Lower reliability at first than with task-specific rubrics -Requires practice to apply well

		rubrics	
Task-Specific	-Description of work refers to the specific content of a particular task (e.g., gives an answer, specifies a conclusion).	-Teachers sometimes using this makes scoring "easier" -Requires less time to achieve inter-rater reliability	-Cannot share with students (would give away answers) -Need to write new rubrics for each task -For open-ended task, good answers not listed in the rubric may be evaluated poorly

In light of the differences shown in Table 1.3, task-specific rubrics seems narrow in comparison to general rubrics as the former is useful only for specific tasks and cannot be generalized. It means they "specify the specific facts, concepts, and/or procedures that students' responses to a task should contain" (Brookhart, 2018, p.1). This may burden users to create a rubric for each new task type which is not that easy. Furthermore, employing task-specific rubrics may not allow this kind to be used at large scales since selecting activities differs from one teacher to another, and choosing comparable tasks is likely to be limited; therefore, using the same rubric is likely to be low as well. This could impact negatively the reliability issue. For overcoming this problem, Jonsson and Svingby (2007) suggested using task-specific rubrics for summative purpose with few levels for the purpose of increasing reliability. Nevertheless, this does not remain a solution, as using fewer levels can flout fairness in the assessment. As it is seen less used for FA, Jonsson and Panadero (2017) see that using few levels in task-specific rubrics could not be appropriate while general rubrics are seen important as they can be shared with students, used for grading and for learning (Brookhart, 2018).

1.6.2.2 Analytic vs. Holistic. Being holistic or analytic is an important way of characterizing rubrics (Brookhart, 2013). This categorization is based on rubrics' content, with analytic rubrics being more thorough than holistic rubrics. Analytic rubrics have separate descriptors, whereas holistic rubrics group the descriptors of each level together (Brookhart, 2013). In fact, some authors see a difference the difference between the two types while others do not.

In the following Table 1.4, Brookhart and Nitko (2008) presented the advantages and disadvantages of holistic vs. analytic rubrics.

Table 1.4 *Holistic vs. Analytic Rubrics. From Brookhart (2013, p. 7)*

Holistic or Analytic: One or Several Judgment			
Type of Rubric	**Definition**	**Advantages**	**Disadvantages**
Analytic	Each criterion (dimension, trait) is evaluated separately.	-Gives diagnostic information to teacher. -Gives formative feedback to students. -Easier to link to instruction than holistic rubrics. -Good for formative assessment; adaptable for summative assessment; if you need an overall score for grading, you can combine the scores.	-Takes more time to score than holistic rubrics. -Takes more time to achieve inter-rater reliability than with holistic rubrics.
Holistic	All Criteria (dimensions, traits) are evaluated simultaneously.	-Scoring is faster than with analytic rubrics. -Requires less time to achieve inter-rater reliability -Good for summative assessment	-Single overall score does not communicate information about what to do to improve. -Not good for formative assessment.

Based on the work by Brookhart and Nitko (2008) and other works (e.g. Nitko, 2001; Park, 2003; Brookhart, 2018). These works were examined by the researcher and simply noted without further development because they are not as well-known as the distinction established by Brookhart and Nitko (2008). On the other hand, mentioning them has the purpose of directing the reader toward them to understand the discussion furthered in this paragraph, we have summarized the comparison between the two types into four major points highlighting the difference between these two types focusing either on the content, function, and use limitations. First, holistic rubrics gather all the aspects in one support such as form with content while the analytic ones give details (Brookhart, 2005). Second, the holistic scoring rubrics provide one score while the analytic ones give sub-scores that are added together to conclude to a final one (Nitko, 2001). It appears that the holistic scoring is more economical than the analytic one due to the details provided about the various aspects of writing for the test taker (Park, 2003). Third, holistic rubrics are mostly destined for activities whose parts are closely linked than separated; therefore, they are more product-oriented (Jackson & Larking, 2012), and less informative (Brookhart, 2005) with no details about the performance, but the analytic ones are process-oriented, diagnostic in purpose (Brookhart, 2005), and detailed in feedback and more informative (Zimmaro, 2004). Thus, analytic rubrics are more time-consuming than the holistic ones (Zimmaro, 2004). Fourth, analytic rubrics are better for classroom purposes; however, for only pure SA, where students cannot see their grades and the information issued from, the holistic ones seem more appropriate (Brookhart, 2013). Furthermore, holistic rubrics are normally used when errors in some part of the process can be tolerated and do not affect the overall assessment (Chase, 1999), and can also probably be used when there is no clear and final correct answer (Nitko, 2001).

Despite the differences made between holistic and analytic rubrics, Wagner and Hibbard (2003) see that both share common functions.

> First, they provide the teacher with a menu of ideas for assessment lists. Second, they help the teacher identify the strengths and weaknesses of the whole class and individual students. This information drives instruction. And third, they help the teacher communicate with other teachers, administrators, students, and parents about the quality of student performance. (p. 30)

The common functions between the two types, as presented in the quotation, do not cancel the difference between them and that difference does not deny that they have common areas to share. Overall, this elucidates the importance of a rubric on its own regardless of its type. In writing, having performance levels with a set of criteria described appropriately offers a forum of discussion to clarify ambiguities at various levels. However for classroom writing aiming at improvement and engaging students properly, the analytic rubrics are more appropriate for the details they offer concerning the different writing features. Analytic rubrics elaborate and clarify the criteria and the sub-criteria which make the idea clearer for students. For example, the criterion of mechanics in writing is not explicit unless it is clearly described as, for example, including spelling mistakes, punctuation and capitalization. In fact, the problem does not lie only in condensing a set of criteria into a major one because the corresponding descriptors to those criteria should also be assembled into a one descriptors. This is what might complicate the issue.

Putting it more clearly, the rubrics used in this study are analytic- instructional and mixed in purpose; general and also task specific. They are analytic in the sense that they provide specifics to the items and include sub-scores, and instructional because they are designed to be at the heart of classroom writing-based assessment. They are general because they are concerned with the characteristics of paragraph writing, and task-specific because they are concerned with the genre under study. More information is provided in the research design chapter (see Sections 2.4.3.1 and 2.4.3.2).

1.6.3 Rubrics' Use

The use of rubrics is discussed from two perspectives: the classical view, for pure grading purpose, and the modern use, for learning purposes.

1.6.3.1 Rubrics' Classical Use. Rubrics' usage is increasingly invading both higher education and schools. They have been the primary choice for many test developers as they allow for reliable assessment of complex performances (Moskal & Leydens, 2000; Jonsson & Svingby, 2007), make teachers confident about their assessment of student writing, enable them to improve the efficiency of grading students' work, and help them to justify the scores assigned (Andrade, 2000). In writing, scoring rubrics could save teachers from an important dilemma which could be summarized in this question: why did you (the teacher) score my (the student) writing in that way? However, scoring rubrics is not as simple as it appears since the problem of grading writing already exists. Teachers' various misconceptions regarding rubrics originate from how the tool is integrated and used effectively as a result of existing grading assumptions (Brookhart, 2013).Research studies show that employing assessment is frequently more beneficial to educators in grading than to students in completing the assessment as educators understand students' needs more than they do (Joseph, Rickett, Northcote, & Christian, 2019).

1.6.3.2 Rubrics' Modern Use. Rubrics have recently been used for formative purposes besides the summative function they have been originally created for (Panadero & Jonsson, 2013). The formative purpose is what we refer to as the modern use while the summative purpose is the traditional one. In their research on student perspectives on rubric-referenced assessment, Andrade and Du (2005) argue that rubrics can serve purposes other than just grading, namely for teaching aims due to their power in clarifying both learning content and outcomes. In this way, rubrics invite both parts to pursue their evidence in the classroom and the school context (Brookhart, 2013). This is to say that the pendulum swings to student centeredness via exploiting rubrics and fostering students to take charge of their own writing and learning in general. The focus has shifted from scores and grades to improving approaches for better obtaining those marks.

To use rubrics in classrooms, Brookhart (2013) emphasizes that they should be well-designed to orient their purposes for both learning and grading. This highlights the outstanding dilemma of implementing rubrics in writing classroom-based AfL purposes where the design would be the inflexion point. This latter could determine the writing rubric effectiveness and foster or undermine students' writing progress. Using writing rubrics for learning targets requires making the outputs of scoring flow in the way of encouraging learning; in other words, leveraging quantitative assessment outcomes to drive learning forward. In particular, Andrade and Boulay (2003) insist that the instructional rubrics should include the following features to support learning: they (a) are written in the language that students can understand, (b) define and describe quality work as concretely as possible, (c) refer to common weaknesses in students' work and indicate how such weaknesses can be avoided, and (d) can be used by students to evaluate their works in progress and thereby guide revision and improvement. As a response to those qualities, this kind of rubrics are seen powerful tools for facilitating self- and PA, especially for aiding them to generate self-and peer-feedback (Jonsson & Svingby, 2007). Those feedbacks when generated might allow students to converse and discuss the assessment outcomes; therefore, rubrics could be used as a support to deliver structured, helpful and effective elaborated feedback as they opt for clear and explicit criteria with well-defined descriptors and quality levels.

In an attempt from Andrade (1997, p. 2), the author summarized the functions of rubrics in six major points. First, “they are powerful tools for both teaching and assessment”. This is to say that rubrics can improve and monitor students’ performance through making teachers’ expectations explicit and paving the way for students to achieve that. The result thus is generally improving students’ quality in learning. Second, “rubrics are useful is that they help students become more thoughtful judges of the quality of their own and others’ work”. This can be explained as follows: through engaging in self- and PA, students may be able to highlight and solve their own and their classmates’ problems; thus, this may increase in them a sense of responsibility. Third, “rubrics reduce the amount of time teachers spend evaluating student work”. This means that by using an appropriate rubric for assessing students’ work, teachers commit to a set of criteria to use each time on every student’s work, saving teachers from having to ponder what to include or omit each time to assess a specific work. In addition to that, the evaluation results of students’ work then can be consistent because the same criteria are used each time to assess every piece. Time-saving may be more noticeable with the holistic type because it is less detailed. Fourth, “rubrics provide students with more informative feedback about their strengths and areas in need of improvement”. This is especially when rubrics are used to intentionally address feedback for learning purposes. Fifth, “teachers appreciate rubrics because their “accordion” nature allows them to accommodate heterogeneous classes”. This may denote that rubrics can be assimilated to an “accordion” as they could meet heterogeneous abilities in the classroom and can shape the aims accordingly. Along the quality levels associated with the descriptors, students could find their real abilities through diagnosing their level to work for further progress. Finally, “rubrics are easy to use and to explain”. This means that rubrics’ content is supposed to be explicit especially when designed appropriately and implemented effectively.

As a final note on the discussion of rubric use, modern does not mean recent and newly created, but it is a matter of revisiting the utility of rubrics to re-exploit them for more authentic and purposive practices. This is in writing classroom-based assessment and for learning purposes to facilitate the writing mission for students and teachers as well. It is more than just grading; it is using the outcomes of grading and the qualitative feedback to nurture the well-being of student progress. The rubric is usually content specific for a specific context. Therefore, for the sake of being more purposive rubrics should not be standardized if we need to make them useful (Turley & Gallagher, 2008). However rubrics can be judged as good as they exceed the limits of assessment and grading and it could be useless if adapting these levels to practice is ineffective. This last point is considered rare to occur in practice (Brookhart, 2013).

1.6.3.3 Rubrics in Writing. Rubrics are crucial especially for assessing performance skills like writing and speaking (Sadler, 2009). They first appeared to earn more validity in grading writing and avoiding subjectivity (Weldon, 2009). Recently, there has been a shift in interest toward the use of co-constructed rubrics as pedagogical tools in which students have a high level of input and investment (Joseph et al., 2019), because teaching writing practices revolve around three major questions: (1) How can students be more fully engaged in the evaluation process? (2) How can we foster more student dialogue and community? (3) How can we increase student writing ownership? (Blumner & Fritz, 1997, p. 234). This demonstrates the important role students play in the classroom, and thus finding ways to involve them appears to be challenging and very critical. One of the ways is exploiting assessment for learning purpose and involving students in the process. This can also be supported by tools such as rubrics to reflect the qualities and the goals to achieve. They are seen important because they help students understand the learning target and the success criteria through coordinating between instruction and assessment (Brookhart, 2013).

In fact, many studies on rubrics have consistently produced positive results, both in the native context, which began in the 1990s, and in the EFL/L2 one, which has recently emerged. In a Chinese context, for example, Wang (2016) found that rubrics were regarded by students as a roadmap, clarifying the highest levels expected in their writing performance and orienting their efforts toward those levels of performance. Sundeen (2014), however, observed that giving the rubric to students or teaching them using the rubric had the same effects on their writing performance. In a broader context, Cho, Schunn, and Wilson (2006) noticed that analytic writing rubrics influenced the writings of 708 undergraduate and graduate students in sixteen courses over three years. Despite having positive effects, Wilson (2007) addressed this question to inquire the reality of rubric use as follows: "Is it possible that discarding the rubric will open the door for some teachers to respond badly or insensitively to student writing? Sure—just as rubrics in the hand of an unthinking, insensitive teacher have predictably unimpressive and harmful results" (p. 66). The author's main point is that the use of rubrics should be approached with caution; they should be carefully developed, designed, and implemented in order to produce positive results.

These findings highlight the significance of using rubrics in writing and the positive impact they have on students' motivation, progress, and ownership. However, more research, particularly in the EFL context, should be encouraged in order to collect more data that can be used to improve the design and implementation of this tool. In agreement with this viewpoint, Jonsson and Svingby (2010) advocated more empirical research on the use of rubrics in writing improvement through explicit instruction, shedding light on the scarcity of such studies.

1.6.4 Validity and Reliability of Rubrics

In this section, we have briefly discussed how validity and reliability are viewed in assessment in general and how it is treated when related to rubrics.

Validity and reliability are also applied onto alternative or authentic assessments (Moskal & Leydens, 2000). In educational context, validity is seen by Messick (1989) as the adequacy, appropriateness, and usefulness of the inferences derived from an assessment system to the purposes of learning. The accuracy of this system is a condition for ensuring its validity. However, validity and reliability issues are still debatable in the literature and scholars still question whether these two concepts should be treated separately, considering reliability part of validity or the other way round. In fact, our concern is not discussing that and deciding which stance is plausible and applicable, but our aim is to focus on their relation to rubrics. In rubrics' research field, the two concepts have been nurtured by various contributions, yet still conflicting. For example, Montgomery (2002) and Jonsson and Svingby (2007) show contradictory views concerning rubric-based assessment. The former claims that increasing rubric reliability may be to the detriment of its validity, whereas the latter see that validity should be prioritized. This is bourgeoned from the work by Moss (1994) entitled "Can There Be Validity without Reliability?" Her answer to that question was "yes". We have to clarify that this work is not directly related to rubrics but rather it deals with assessment overall. Moss's position is based on the works of the psychometricians Cronbach and Messick (1989), who argued that validity is not related to the assessment tool itself but it is linked to individuals' decisions on the basis of that tool. This is a brief summary of the discussion held about the two concepts, each alone and the relationship between them, but the field is still open to further debates. In this context, McConnell (2013) sees that research about validity has a narrow vision.

To achieve validity and reliability in using rubrics some few suggestions have been proposed. Starting with validity, Andrade (2005) claims that a rubric, especially the instructional one, to be valid should meet "reasonable and respectable standards and with the curriculum being taught" (p.29). For Reddy and Andrade (2010), language clarity and appropriateness are central concerns in scoring rubrics' validity. On the other hand, to be reliable rubrics should result in similar ratings when used by different people (Andrade, 2005; McConnell, 2013). As a result, they should be well-developed, focusing on training raters and care about calibrating sessions (Jonsson & Svingby, 2007; Reddy & Andrade, 2010). Put together, Rezaei and Lovorn (2010) questioned the reliability and validity of rubrics used for assessment purposes by indicating that raters with and without rubrics are concerned with the mechanics of writing rather than the content. We can see that rubrics' validity and reliability is not much developed or fully decided upon when directing their concern toward learning purposes.

1.7 Self- and Peer Assessment

Self-assessment and PA are reported as recent new forms of assessment, but they have actually been found in some areas of education (Topping, 2003). Many researchers find that using both self- and PA engages students to be active participants in the classroom, via identifying and assimilating their own and their peers' work qualities (e.g. Hyland & Hyland, 2006; Lundstrom & Baker, 2009). In line with this, Black and Wiliam (2009) argue that self- and PA in addition to classroom questioning, sharing success criteria, commenting, and summative and formative marking, should be integral parts in the classroom. Primarily, self-and PA are often interrelated and go hand in hand (Fautley & Savage, 2008) but certainly they have intricate relationship because they share common aspects and each contributes in promoting the other.

To review self- and PA in this thesis, we prefer discussing their definitions and importance each separately despite being "typically combined together" (Falchikov, 2005, p. 120), then shedding light on the relationship between them, and finally discussing their implementation and the factors affecting their occurrence and effectiveness.

1.7.1. Self-assessment: Definition and Importance.

The interest in self-assessment has been one of the main areas of research in contemporary education and educational psychology research since the seminal review made by Falchikov and Boud in the late 1980's (Boud & Falchikov, 1989; Falchikov & Boud, 1989). The increasing interest in the role of self-assessment in language learning and teaching is a logical outcome of increased interest in learner-centered teaching and self-directed language learning (Peirce, Swain, & Hart, 1993). For example, Ben-David (1999) asserted that performance assessment requires feedback, explicit criteria and self-assessment. So self-assessment is a term that can vividly animate the importance of involving students in classroom.

Before defining self-assessment as a concept, we have noticed that the definitions available in the literature can be divided into two major orientations, based on the purpose for which it is employed. The first orientation, which emerged first, is self-assessment for summative or pure grading purposes, whereas the second is for learning purposes. In this section, our focus is primarily on self-assessment for learning purposes and any inclusion of summative self-assessment requires that this directly or indirectly contributes to learning progress and promotion because summative testing outcomes have been discussed to be flawing in AfL (see Section 1.1)

Initially, self-assessment was viewed as students guessing or predicting their grade, examining their correlation to their teachers' ones (Falchikov & Boud, 1989), but recently the view has appeared to have shifted from self-grading with a summative purpose to self-feedback for enhancing learning (Panadero, Lipnevich, & Broadbent, 2019). In line with this, Falchikov (2005) argues that self- and peer grading, where students grade their own works or that of their peers using marking schemes or using model answers, are seen a special case of self- and PA. However, self-assessment for summative purposes is unfortunately still pervading and focusing on predicting students' grades and their correlation to teachers' scores despite shifting toward exploiting self-assessment for promoting learning (Panadero, Lipnevich, & Broadbent, 2019). In her review, Andrade (2018) found that thirty out of fifty-two publications published on self-assessment from 2013 to 2016 focused on accuracy. This is to highlight that despite the recent shift in focus to self-assessment for learning purposes, the literature on self-assessment is still overburdened with studies on self-assessment for only summative purposes.

Regardless of letting students grade their own works, self-assessment can be viewed as an opportunity offered to students to reflect on the quality of their work in order "to learn more, make improvements and perhaps even earn a higher grade...to collect information about their own performance and see how it matches their goals and/or the criteria for their work" (Andrade & Du, 2007, pp. 159-160). Reflecting on their works, students should be given a chance in self-assessment not only to revise their work but to question their pitfalls and look for solutions. In relation to that, self-assessment could potentially be a powerful classroom practice, especially with productive performances like writing. On the other hand, Andrade (2018) defines self-assessment through focusing on its purpose as it can "generate feedback which promotes learning and improvements in performance" (p. 377). This can denote that self-assessment is to orient students toward engaging in classrooms as active participants and foster in them the will to reflect upon their "own learning processes, styles and outcomes" (Topping, 2019, p. 14) through generating feedback which might come from judging their own progress and being responsible for it. As a result, self-assessment can assist students in setting their objectives and understanding that they must study for themselves and not for anybody else, such as their teachers or even their parents (Hanrahan & Isaacs, 2001). In this context, self-assessment is often seen as "a continuous longitudinal process that activates and integrates the learner's prior knowledge and reveals developmental pathways in learning" (Topping, 2019, p. 14) to especially trigger students' critical thinking and promote their further cognitive and meta-cognitive potentials. For example, Jonsson and Botella (2017), in their meta-analysis, found that the role of self-assessment interventions is to promote students' use of learning strategies and its effects on motivational variables such as self-efficacy. Other meta-analyses also revealed that self-assessment impacts on student achievement (Brown & Harris, 2013) self-regulated learning and self-efficacy (Panadero, Jonsson, & Botella, 2017). In short, self-assessment has the potential to influence various aspects of learning, going beyond merely surface self-grading.

Some researchers concentrate on self-assessment purpose while others focus on the performance itself, and yet others on the implementation method. In this vein, Andrade and Du (2007) see that there is no standard definition of self-assessment, whereas Panadero, Lipnevich, and Broadbent (2019) contend that the common underlying idea in self-assessment definitions focuses on learner engagement with the process or product of their own learning to describe how they perceived their progress or result, whether for summative (self-grading) or learning targets. Andrade and Heritage (2018) and Panadero, Brown, and Strijbos (2016), particularly, suggest that future research about self-assessment will operationalize this concept equating it with self-feedback and encouraging its use formatively. According to Stiggins (2007), the significance of self-assessment lies in providing students and teachers with constant access to detailed feedback, that is perceived not overwhelming, to enable students design their path toward specific achievements highlighted by their teachers. As self-assessment is based on self-feedback, which in turn modifies processes and products to improve learning, it is regarded as an important component of AfL because it can provide a space for feedback with opportunities for students to reflect on and improve their work as well as providing teachers with necessary information to modify their instruction accordingly (Panadero, Lipnevich, & Broadbent, 2019). Students may perceive self-assessment effectiveness when they are aware of their teacher's expectations; they may be helped by using self-assessment to check their work and guide revision; and thus they may conclude that the benefits of self-assessment can lead to improvements in grades and work quality, as well as further motivation and learning progress. As a result, self-assessment can involve students and teachers in a cycle of feedback/reflection under teacher scaffolding and guidance before, during, and after performing. This implies that tracking self-assessment performance should not be restricted to a certain period in performance because all three stages- before, through, and after- appear to be interwoven and complementary. In this regard, Andrade and Du (2007) found that students gain positive attitudes toward self-assessment, especially after extended practice. This idea is extensively discussed in Section 1.7.5 about factors affecting self- and PA effectiveness

To conclude this section, we can say that one of the main focuses of this study is on self-assessment that promotes students' writing and improvement within AfL. In this vein, self-assessment should empower students' abilities to diagnose, reflect, evaluate, and assess their own writing composing strategies in purpose of knowing their strengths and weaknesses to exploit the former and work on the latter for the aim of improving their writing. In fact, this goal should not only be limited to the short term goals, but that should be extended to further motivation (Andrade & Du, 2007) and cognitive abilities since it is supported by theories of constructivism and learner autonomy (Chen, 2008) to help students be lifelong learners and self-regulate their learning (Andrade &Valtcheva, 2009). In this context, Boud (2000), for instance, argues that self-assessment helps students to rework throughout their life to face new challenges and Falchikov (2005) suggests that being a lifelong learner means being lifelong assessor. Additionally, implementing self-assessment for learning ends might prevent students from believing that the teacher is the ultimate authority and alert them to the need to rely on other learning aids in order to be self-reliant and autonomous (Taras, 2008).

1.7.2 Peer Assessment: Definition and Importance

To a widespread prominence, PA came about 20 years ago and has widely become popular in the last ten years (Topping, 2019). In fact, PA started in higher education and then disseminated to primary, secondary and vocational education. It notably started in the writing field, especially the academic kind and writing as a foreign/second language in activities such as student presentations and then emerged in other fields such as science and mathematics (Strijbos & Wichmann, 2018). The review of this concept has marked at the start with the work of Topping (1998) for purposes of filling the gap, providing typologies, and setting theoretical underpinnings of PA despite the existing reviews about self- assessment that was presented by Boud and Falchikov (1989) and Falchikov and Boud (1989).

The word PA is sometimes used interchangeably with other phrases such as peer review, peer feedback, and peer response; however, the focus of this study is on PA per se, omitting any substitution because the latter has lately been criticized. If any of the above concepts appears throughout this thesis, it is because the idea is tightly related to PA, defying any assumption of equivalence. Topping (2019) shows that many link peer review and peer feedback with PA whereas this is not the case since the former is used to examine research works and occasionally with school children, while the latter is not deemed PA if it is not detailed or reciprocal.

Despite the various definitions of PA regarding its goals and outcomes, this concept is often used as a simple stylistic variable (Lundstrom & Baker, 2009). As a matter of fact, PA was viewed at the beginning of its emergence, like self-assessment, as peer grading/scoring but recently it has shifted to be for learning purposes. This review will exclude any definitions related to pure summative PA, focusing specifically on PA for learning purposes. This latter focuses mostly on delivering detailed feedback to assessees for the sake of amending their work (Topping, 2019). PA for learning can assist to broaden the scope of peer feedback and provide a variety of feedback types to fit student understanding. This comprises grading, spoken and written comments that foster communication and involvement in order to achieve the teamwork that PA seeks (Strijbos & Wichmann, 2018). To define this concept, Topping (2019), the pioneer in this field, sees that PA is

> an arrangement for learners to consider and specify the level, value, or quality of a product or performance of other equal-status learners, then learn further by giving elaborated feedback to and discussing their appraisals with those who were assessed to achieve a negotiated agreed outcome. (Topping, 2019, p.1)

This is to say that PA contributes to creating an interactive and effective learning atmosphere, where students give and receive feedback and respond to either act accordingly and constructively. While peer assessing their works, students are engaged in meaningful student interaction where they are supposed to evaluate and discuss the criteria related to their assignment, and by so doing they are able to judge their own work and identify their strengths and weaknesses (Lundstrom & Baker, 2009; Topping, 2019). In addition, they can target areas for “remedial action, and develop metacognitive and other personal and professional skills” (Topping, 2019, p. 3). This denotes that PA has a wide range of positive outcomes that could be achieved if appropriately implemented.

As feedback is seen primordial in PA, Strijbos and Wichmann (2018) try to link the two concepts together, proposing the following definition:

> Peer assessment and peer feedback as an interpersonal and collaborative practice refers to a process in which one or multiple assessors (individual student or group) provide quantitative (scores or grades) and/or qualitative (written and/or oral comments) feedback on the product of one or multiple assessees (individual student or group), who subsequently interprets the feedback and applies it (partially or fully) to revise the product and/or interacts with the assessor to clarify their interpretation and/or justify the extent to which it was applied. (p.3)

For the purpose of improving and promoting learning, PA can exploit peer feedback provided to establish a relationship between the sender and the receiver. In this mechanism, students could be acting as instructional source for each other (Black & Wiliam, 2009) and collaborate among them to share experience in a mutual way. Within PA practice and from scoring perspective, Bostock (2000) sees that PA is the assessment of learners to their peers providing them with both formative and summative grading. In fact, the idea emphasized is that scoring could be part of FA, where feedback takes part, as scores generated by PA could also inform, motivate, or document students’ further progress.

To summarize, PA can be viewed as an instructional activity that engages students in a learning entity to collaborate and feedback each other through providing constructive and elaborated feedback. This latter could be supported by scores, if needed, to inform the receivers about their progress and allow them to reflect on their works for further promotion. In this context, PA is not equated to peer feedback despite the equivalence pervaded in studies dated back to 2019. A last point to shed light on is that scholars have recently debated PA benefit, questioning which students’ levels can be influenced by PA/ low achievers, high achievers, or both. Initially, it was stated that PA helped all students without specifying which level, but some objections have lately arisen. For instance, in a study by Li and Gao (2016) the researchers found that PA is supposed to positively affect low achievers over the high achievers while Williams, Carroll, and Hautau (2005) find that high achievers show decrease in their abilities when being engaged in PA. These assumptions are actually reported from studies about PA out of the writing field; hence, further studies are needed in this field to support or reject any of the assumptions formulated.

1.7.3 Relationship between Self- and Peer Assessment

In higher education, self- and PA are frequently regarded as a single assessment innovation. This sparked issues about whether they have distinct roles or go hand in hand, whether they are similar or different, and what sort of connection they have (Falchikov, 2007). Briefly, if someone asks the following question: is there a relationship between self- and PA? The answer is definitely yes. Despite being separate in terms of performance occurrence, self- and PA are potentially interrelated as they both require student judgments but those judgments are different (Adachi, Tai, & Dawson, 2017). PA is sometimes considered as a crucial supplement to self-assessment and may come before it (Black et al., 2003). Learners may not take their own assessment seriously; therefore, PA might come to teach them that assessment exists and is important to consider. In this regard, PA might provide feedback in the way students can understand and can also teach them how to self-assess their works (Bostock, 2000; Black et al., 2003). On the other hand, Dichkut (2003) finds that combining self- and PA encourages knowledge and creativity, especially higher order thinking skills. The meta-cognitive ability related to combining self- and PA is related to various skills like interpreting, problem solving, evaluating, reflecting (creating internal feedback), and evaluating their level of understanding; “what one knows about one’s own cognitive ability” (Amhag, 2013, p. 94). In a study carried out by Falchikov (1986) looking at students’ views about self- and PA, the students felt that both assessments made them think, learn, be critical and be structured.

1.7.4 Implementing Self- and Peer Assessment

Despite the fact that they are legal procedures in classrooms, self- and PA are not really welcomed even among those who take assessment seriously because their implementation is reported to be quite challenging (Fautley & Savage, 2008). In this section we discuss the implementation related to introducing student assessment, self- and PA, for boosting learning and not for pure summative purposes. For so doing, we discuss self- and PA implementation, going through the overall guidelines, conditions, procedures and characteristics related to the implementation and then narrowing the interest down to the main factors for effective implementation

1.7.4.1 Guidelines, Conditions, Procedures and Characteristics. Given that self- and PA are both student evaluation performances, it would be fair to combine a set of points to cover the overall implementation of both assessments. These suggestions are based on Wyatt-Smith and Klenowski’s work (2014). The authors basically advise scaffolding and supplying information to include students in self- or PA in writing activity. They then suggest the following recommendations:

1. providing a comprehensive outline of what is involved in the completion of the assessment task,
2. planning materials to support the student in completing the assessment (particularly when it is a rich or extended task with several components),
3. opting for an annotated commentary of the assessment task to scaffold the student’s response,

4. exposing an exemplar of a completed task and an annotated version of the exemplar to highlight how the key requirements of the task have been addressed, and
5. establishing specifications (sometimes called a rubric or criteria sheet) that outline the criteria and standards by which the task will be assessed (p. 201).

In addition, we synthesized a set of characteristics and conditions about self-assessment to be taken into consideration while implementing it based on the works presented by Boud (1991), Goodrich (1996), Stalling and Tascione (1996), Adams (1998), Hart (1999), Gregory et al. (2000), Andrade and Du (2007), and Andrade and Valtcheva (2009). They are the following:

- raising awareness about self-assessment (Andrade, 1996; Stalling &Tascione, 1996; Andrade &Valtcheva, 2009)
- involving students in identifying the assessment criteria (Boud, 1991)
- sharing assessment criteria (Boud, 1991; Andrade, 1996; Andrade & Du, 2007; Andrade & Valtcheva, 2009)
- scaffolding self-assessment occurrence through direct instructions and assistance (Stalling & Tascione, 1996; Andrade, 1996; Andrade & Valtcheva, 2009)
- exposing students to practice (Andrade, 1996; Andrade & Valtcheva, 2009)
- following self-assessment occurrence to check its appropriateness and giving opportunities for improvements (Andrade, 1996; Andrade & Valtcheva, 2009)
- keeping self-assessment formative (no self-grading) (Andrade, 1996; Andrade& Valtcheva, 2009)
- using work models (Andrade, 1996; Hart, 1999; Gregory et al., 2000; Andrade & Valtcheva, 2009), and
- generating feedback that is used in guiding student revision (Adams, 1998; Andrade & Du, 2007).

Despite being specific to self-assessment, these characteristics and conditions can also be applied to PA. Basically, implementing self-and PA should be contextualized, integrated, and holistic. This requires that students need instructions, opportunities, practice, and feedback to assess the progress they made through creating a discussion forum where they expose their issues and inquiries related to self- and PA, such as how using the criteria and how applying those criteria. This might clear up numerous uncertainties and ensure a better performance. In fact, all the aforementioned points cannot be applied to all scenarios since the status quo varies in terms of students' skills, learning perceptions, and assessment eclecticism.

1.7.4.2 Factors Affecting Self- and Peer Assessment Implementation. Digging out the literature about self- and PA implementation has alerted us to the importance of shedding light on major factors which might strongly influence the implementation. They are involving teachers, involving students, training students, and using assessment tools. In what follows, we shall examine how that could influence the overall implementation.

a.Involving Teachers. In fact, implementing self- and PA is not linked with a set of fixed procedures or guidelines to follow as it is more linked with students' full understanding of this performance and its importance to their learning. This requires that teachers should first understand this performance to be able to transfer it to students. To empower its occurrence and make it instructional and a learning activity, Boud (1999) suggested involving external sources of feedback such as teachers or peers. Regarding this idea, some scholars suggested shifting from total student control (Edwards& Sutton, 1991) to pre-dominantly teacher control (Taras, 1999). Consequently, before moving to practicing self- and PA, metacognitive teaching about the processes should be prior, as Brooks (2002) noted:

> Students must be trained in the required metacognitive skills by teachers modelling the processes for them. Teachers can do this by sharing marking exercises with classes, using exemplification material to show how criteria are applied and how judgements are reached. (Brooks, 2002, p. 70, as cited in Fautley & Savage, 2008, p. 52)

Creating a space of discussion between teachers and students would enable increase awareness in relation to and with their learning. When it comes to debating teacher intervention in managing self-and PA, it is plausible to see that the teacher has a role to perform, either a minor or a major one, in introducing those performances because there are some ambiguities to clarify and guidelines to explain even if students were already familiar with those practices. For further clarifications, self- and PA are intrinsically linked with a set of goals and used for a specific learning context, i.e. for a given course most of the time; thus, variations in goals and learning context would generate some confusion while approaching student assessment performances. In this context, Spiller (2012) emphasizes that the conversation should be intensive with students before introducing self-assessment. Furthermore, Topping and Ehly (2001) suggest organizing, managing and monitoring PA by teachers to obtain effective outcomes. This may help students understand student assessments processes and their goals to gain readiness and challenge the occurrence. In writing, for example, involving teacher and peer collaboration during teaching writing helps in negotiating meaning (Reid, 1994) because teacher guidance helps students collaborate and provide feedback to each other, face the responsibility toward discussing, promoting each other's ideas, linking their progress to their personal aims, and reflecting on their own learning and experience the achievement feeling.

b.Involving Students. Before debating student inclusion or not and criteria negotiation , researchers are discussing whether or not using a set of criteria in self- and PA. Because we are concerned with using the criteria in self- and PA in this study, we avoided discussing non-use of the criteria, and we only pointed to it to provide a complete picture for the reader regarding this point-This might generate further axis of research.

In some cases, merely sharing success criteria is not sufficient, and further additional efforts should be added. For example, James (2013) suggests exposing students to materials such as samples of assessed works and works with mistakes and shortcomings. Providing samples could be the simplest way to make students understand how a given written work should be, what pitfalls to avoid and what strengths to emphasize on. To discuss student involvement in self- and PA, we have dealt with each performance alone because no works found discussing student involvement in the two performances gathered in one study. This appears logical as student role is different i.e. in self-assessment students are dealing with themselves while in PA students are in a position to interact with their peers.

- **Involving Students in Self-assessment.** The difficulty of engaging students in self-assessment following Broadfoot, James, Mcmeeking, Nuttall, and Stierer (1988) stems from two factors: the students' unfamiliarity with it, and the ambiguity, unavailability, and misperception of the assessment criteria, which were usually norm-referenced. On the other hand, McNamara and Deane (1995) suggested that it is up to students to choose criteria and guide self-assessment for learning. Therefore, it appears that involving students in the selection and development of the assessment criteria is essential (Boud, 1991/1995; Race, 2001; Taras, 2008; Topping, 2019). It was argued that students might become good judges of their own strengths and shortcomings and establish their realistic objectives by including them in the identification and formulation of the evaluation criteria (Oscarson, 1997, as cited in Brindley, 2001). As a result, self-assessment should be viewed as a contextualized and embedded activity in a particular learning environment (Panadero, Lipnevich, & Broadbent, 2019) rather than an isolated and individualized task.

To engage students in establishing criteria, Taras (2008) suggests that: (1) students create their own criteria, (2) students negotiate them with their teacher, or (3) students negotiate the criteria with their teacher in the context where they come from. On the other hand, Leach (2012) found that no student opted to dispute the criteria provided while performing self-assessment without explaining what reasons were behind that reaction. This could be explained in several ways: (1) the students might be completely satisfied with the criteria, (2) the students might not understand the importance of performing self-assessment, (3) the students might not know what criteria were and their importance for self-assessment, and/or (4) the students may not have been given a chance to negotiate the criteria using various techniques such as classroom conferencing and oral or written feedback because the author did not elicit what the students wanted.

- **Involving Students in Peer Assessment.** To conduct PA independently students must understand the learning objectives (Black et al., 2003; Harlen, 2005) as well as the evaluation criteria (Harlen, 2005; Sebba, et al., 2008). The extent to which the assessment criteria are understood can have an impact on the quality and accuracy of the feedback provided (Black & Wiliam, 2009). Only then will students be able to communicate with their peers and offer feedback (Sebba et al., 2008). Discussion with a peer about a project might raise awareness of hidden opportunities or challenges (James, 2013). This may also prepare students to be responsive toward their counterparts' reactions.

Discussion in PA appears to be critical prior to, during, and after doing it. Prior to performing PA, incorporating students in the discussion helps primarily in clarifying the task, and discussing during-performing helps in reflecting and adjusting, if any, the performance. However post-performance discussion can help in responding to the assessment outcomes. This is explained by the importance of engagement among students, which may increase social communication, problem-solving abilities, work and time management, and students' selection of acceptable assignments to do. Engaging students in PA may make them a source of assistance by revealing their opinions, thoughts, and suggestions. If they do not directly influence their peers' actions, their voice will be heard somewhere else in the ultimate result.

Another critical point to emphasize when engaging students in PA is procedures the assessors follow (Topping, 2019). These procedures vary in degree; they may include strict and instruction guidelines to follow by both the assessor and the assessee at the beginning, and later students may subsequently learn to accept responsibility and create more feedback in their own way. Putting students under strict guidelines could help as it could reduce or inhibit their motivation and their overall progress. As a result, this might lead to less awareness frustration and even more anxiety. These guidelines, conditions, procedures and characteristics are fully discussed in Section 1.7.4.1 above.

c.Training Students. The issue of whether or not students should be trained is usually highlighted in using summative self- and PA assessments. When these performances are intended for learning purposes, this topic receives little attention, and the majority of the discussion revolves about scoring assessment. For example, Hanrahan and Isaacs (2001) found that training students raises the reliability and the validity of the assessment performance, and Fontana and Fernandes (1994) found in their study that the regular use of self-assessment in mathematics, after training sessions, with children of primary school revealed improvements in scores compared to the control group. On the other hand, in a study by Ninness, Ellis, and Ninness (1999), the students showed higher motivation and grades and acquired lasting math performance after being trained, but Mowl and Pain (1995) indicated that preparing students to perform self- and PA is helpful but not necessary.

In light of self- and PA for summative grading purposes, and the lack of works discussing this issue when it comes to orienting self- and PA for learning purposes, we believe that training students in this context could be done or not depending on the underlined objectives, the students, the teacher, and the overall classroom environment. For example, if the goal is to examine students' awareness and/or reactions to self- and/or PA for the first time, collecting students' opinions or examining their backgrounds surrounding those performances may be preferable to teaching students. However, for the purpose of helping students to get benefits from those student assessment performances, training them could be useful and helpful.

d.Using Assessment Tools. Using assessment tools in doing self- and PA is another debatable issue. For example, using rubrics which display a set of explicit criteria upon which students can rely to evaluate their works goes in line with self- and PA characterization and practicality. In this context, Panadero et al. (2013) have found that when PA is associated with rubrics, undergraduate students' writing abilities are improved. In addition to that, it is found that combining classroom-based performance assessment with use of criteria and standards is highly enabling of self-assessment and self-regulation of learning because sharing "meaning-making of criteria and standards among teachers and students is a socially interactive process" (Colbert & Cumming, 2014, p. 214). Sharing the criteria could be, to some extent, a helpful way to perform self- and PA. This is to say that the criteria and descriptors presented by rubrics could give an idea about student works' evolution and that might also offer a kind of progress scale against which students could compare their improvement. Finally, in their meta-analysis Panadero, Jonsson, and Botella (2017) note that self-assessment can happen in its most rudimentary form without any aids as it can include, more wide-ranging and supporting tools, including instruments such as criteria, prompts, and rubrics.

1.7.5 Factors Affecting Self- and Peer Assessment Effectiveness

Student assessment has been shown to be important, but it is viewed in various ways: being taken for granted or addressed explicitly (Black & Wiliam, 1998) and being either integral to classroom learning or introduced as a separate activity. Despite being seen a beneficial performance, self- and PA are perceived as difficult, challenging, and time demanding (Hanrahan & Isaacs, 2001). In this context, there are several factors that might influence self- and PA efficacy. When researchers discuss their effectiveness, they mostly link them to summative goals. In this section, we will look at each performance individually focusing on the factors that influence its effectiveness. This is because no work was found dealing with the two performances concurrently, or discussing them in connection to learning objectives.

1.7.5.1 Factors Affecting Self-assessment Effectiveness. Comparing the factors affecting validity and reliability of self-assessment to TA which seems to be “a little lower and more variable, with tendency to over-estimation” (Topping, 2019, p. 15), Topping proposed a list of factors in 2003 and another one in 2019 explaining this difference. In 2003, the author explained the difference in correlation between self- and TA, for summative purpose, which ranged from 0.40 to 0.94. She suggested factors such as experience, age, subject, race, and gender difference to be the reasons of this difference. In 2019, the same author suggested additional factors:

> the ability of the learner, the amount of scaffolding, practice and feedback and the degree of advancement in the course, rather than chronological age…the nature of the product or performance assessed, the nature and clarity of the assessment criteria, and the nature of assessment instrumentation. (Topping, 2019, p. 15-16)

The factors proposed in 2019 are explicitly pointing to the difficulty of the performance when used for purposes other than just pure grading. In other words, the author has clearly suggested a list of factors affecting self-assessment effectiveness for learning ends and conveys that self-assessment had emerged to be for pure scoring purpose, following the list of 2003, but shifted to learning goals, following the list of 2019.

a.Self-assessment and Time. Being time-saving when emphasizing its final goal (Looney, 2008) and time consuming when emphasizing its performance (Hanrahan & Isaacs, 2001), using self-assessment for the first time seems problematic. Cowan (1988) reported that his first experience with self-assessment was disastrous because students were unable to know and how to start in addition to his lack of knowledge about how to facilitate autonomous learning and self-assessment. This might also be worse when using self-assessment for the second or third time because the impact of the prior experiences might prevent or slow down students to adjust their vision toward self-assessment, especially when they failed in the prior experiences. This could also influence teacher experience with self-assessment and could shake their confidence toward incarnating this performance in classrooms.

b.Self-asssessment and gender and Ethnic Group. Following Gipps (1994), this difference concerns the process itself and negotiations with teachers. In this context, the same author found that while males tended to challenge teacher’s assessment with a keen sense, girls were ready to enter in further discussion and negotiation. This could be explained by the males’ characters of being direct and avoiding discussions and the reverse for girls. However, in a study by Andrade and Boulay (2003), the researchers did not report any gender difference in scoring using self-assessment, but they did when reacting to feedback.

c.Self-assessment and Student Responsiveness. Students’ responsiveness to self-assessment is another factor that can impact the effectiveness of this type of assessment. Falchikov and Boud (1989) noticed that low achievers, beginners or students with rudimentary skill, tended to overestimate their abilities while their counterparts, high achievers or students with advanced skills, opted to underestimate their abilities. This reaction might be explained differently: (1) low achiever students would neither understand nor perceive accurately the importance of self-assessment,

(2) they would not be peaceful with their weaknesses and would take their mistakes to be illegitimate, and (3) they would be unwilling to spend efforts to improve their learning, or they would be making efforts without showing that explicitly but they still defend their work to be the best until they fix their problems. This latter can be related to lacking self-confidence toward accepting their problems and working for further improvement.

1.7.5.2 Factors Affecting Peer Assessment Effectiveness. PA can be of paramount importance, but there are factors that prevent this assessment practice to achieve its potential. In what follows, we discuss the major factors that may interfere with PA effectiveness.

a. Peer Assessment and Self-confidence. When introduced in classroom, PA can create in students' lack of confidence both in their own and their peers' abilities to assess (Ballantyne, Hughes, & Mylonas, 2002). This could be raised due to their apprehension to accurately undertake such kind of assessment. In the EFL context, for example, this could also be explained by a lack of linguistic capacities students may suffer from especially lacking accuracy and fluency in language use. Consequently, opting for a clear language instruction and explicit criteria such as rubrics would be helpful.

b. Peer Assessment and Scaffolding. Scaffolding PA and teacher guidance is assumed to be important. In this context, Lave and Wenger (1991) explained scaffolding to be the zone of proximal development which has three discrete interpretations:

1. Zone of proximal development is seen in the distinction made between the initial performance with support and assistance or collaboration of more experienced people, and the subsequent performances without assistance.
2. The second interpretation is that zone of proximal development is cultural. It is explained by the cultural knowledge provided by socio-historical content and a person's everyday experience.

3. From a collectivist or societal perspective, the third explanation, zone of proximal development is the distance between the everyday actions of a person and a collectively generated solution to a social activity's historical new form.

PA could go in agreement with the first interpretation because students respond to the feedback received, especially when peer assessors have higher abilities compared to assessees, to reflect upon the initial situation.

c. Peer Assessment and Time. Time is an important factor that could affect PA, and being aware about the appropriate period given to this activity is a must. The amount of time given would definitely affect this activity either positively or negatively; giving less time when more is needed would prevent students from achieving effectiveness and giving more when less is needed may bore and slow down students' progress. With enough time and even with less skill, the assessor in PA can produce assessment of equal reliability and validity to the one of teachers (Topping, 2019).

d. Peer Assessment and Cultural Background. Cultural background is an important factor regarding responsiveness to PA. For example, Topping (2019) found that students from the Middle East show resistance and this is primarily linked to gender as boys might be quite reluctant to accept girls' advice, but resistance of the Asian students is explained as follows: the more various the answers provided by students the more the resistance is (Topping, 2019). Concerning this issue of gender resistance, opting for APA would be the appropriate solution to reduce it while for resistance generated by various answers, using a template or a guided feedback such as introducing rubrics designed for learning purposes or even using prompts and checklists would solve the problem.

e. Peer Assessment and Social Relationship. Establishing social relationships prior to learning is very important. The dialogue between less and more knowledgeable persons helps in developing the inter-mental knowledge in the zone of proximal development of the less knowledgeable minds (Vygotsky, 1987). The relation between teachers and students should be given more importance. Marking a significant distance between teachers and students would lead to students' loss of control over their progress because they might feel neither supported nor guided especially when they need that. Overall, social interactivity between teachers and students and between students and their peers would impact learning progress.

1.7.6 Conclusion

Self- and PA have recently been viewed as student assessment practices that empower AfL. They go beyond merely grading and scoring and flow in the AfL implementation purposes; improving, progressing, and promoting learning outcomes. Self- and PA could be seen similar in terms of occurrence and in terms of the many benefits they may offer whereas the difference lies within feedback delivery and its transmission. These differences could be summarized in the following questions: How do students deliver feedback in self- and PA? How do students perceive their peer feedback? How do students perceive their own feedback? These could be the major differences between these two concepts, but the remaining aspects could be seen very similar, namely the implementation procedures, guidelines, and the factors affecting their occurrence. Another point worth focus is validity and reliability of self- and PA. These two difficulties are mostly associated with summative self- and PA. To the researcher’s knowledge, these two issues have not yet been considered in the context of formative self- and peer feedback, which is concerned with quality and overall growth and improvement rather than accuracy and quantity.

1.8 Portfolio Assessment in Writing

A portfolio is a tool that can be added to both assessment (Wiggins, 1989) and constructivist learning (Zimmerman & Schunk, 2001). The idea of portfolio assessment seems to have started with Belanoff and Elbow (1986) who demonstrated that this tool offers a range of benefits to students, teachers, and program administrators. Dating back ten years ago or more, portfolios were used in many contexts for a variety of purposes (Hamp-Lyons & Reed, 1990; Wolf, 1989). Their emergence is linked to the field of visual arts and performance-based disciplines like architecture, design, photography, and other fine arts classes in high schools to move later to the writing field and many other disciplines (Weigle, 2002; Davies & LeMahieu, 2003).

In foreign language instruction, portfolio use emerged to assess both culture and language learning (Allen, 2004). Portfolio assessment, like other movements in writing assessment, was first used in L1 contexts before spreading to L2 settings and most of the literature available on portfolio assessment comes from L1contexts (Hamp-Lyons & Condon, 2000; Weigle, 2002). To shed light on its first appearance, portfolios have been used in the British educational system for over 50 years and by teachers and educators of the United States in 1970s. In writing, using portfolios to assess students' writing performance has emerged in L1 context since the mid-1980s (Belanoff & Dickson, 1991) and then extended to the L2 contexts later on. Therefore, according to Lam (2015), its utility to sponsor classroom-based assessment, especially in writing, is under represented in EFL settings (Lam, 2015).

In this section, we primarily focus on what a portfolio is and what qualifies a portfolio as such. Then, we present the various types, with a focus on the working portfolio as our main concern, and discuss how to treat portfolios and when portfolios become for AfL to provide strong evidence for learning progress. Finally, we discuss portfolio validity and reliability.

1.8.1 What Is Portfolio?

Writing portfolio can be defined as folders or websites (i.e. electronic portfolios) which include a range of evidences about student learning that documents their growth in writing through active self-reflection (Genesee &Upshur, 1996), or it is a purposeful collection of student texts that demonstrates students' effort, progress, and achievement in writing over a period of time (Weigle, 2002). From the angle that visualizes assessment as a way toward ongoing learning and fostering the continuity of assessment to capture a rich array of students' knowledge, abilities and competencies in realistic contexts, Arter and Spandel (1992) adapted a definition of portfolios presented by Northwest Education Association as follows:

> A student portfolio is a purposeful collection of student work that tells the story of the student's efforts, progress, or achievement in (a) given area(s). This collection must include student participation in selection of portfolio content; the guidelines for selection; the criteria for judging merit; and evidence of student self-reflection. (p. 36)

This definition places the portfolio in the forefront as an essential medium for creating a space of communication by engaging students to be active participants in reflecting on their work for purpose of progress. Thus, portfolios can be envisioned as "learning tools to understand and monitor development, as well as a non-mechanistic, outcome-oriented collection of evidence on effort" (Smith & Tellima, 2003, p. 626). In fact, the evidence collected from portfolios demonstrate students' growth, progress, and achievement and that can be qualitative, quantitative (for pure grading purposes), or both. The achievement to reach, whatever the type is, is generally determined by the purposes or goals the portfolio is used for.

Despite being defined differently, some pre-requisites must be taken into account in order for a given portfolio to successfully gain its name. The first pillar is that only collecting students' works does not make a portfolio, reflecting (second pillar) and selecting (third pillar) are essential aspects to make a given folder called as such (Weigle, 2002). On the other hand, a portfolio, then, is a portfolio when "it provides a complex and compehensive view of student performance in context", "the student is a partcipant in, rather than the object of, assessment", and it "provides a forum that encourages students to develop the abilities needed to become independent, self-directed learners" (Paulson, Paulson, & Meyer, 1991, p. 61). This is to say that portfolio's main purpose helps in activating and not merely reporting about the learning path.

Even though it was initially developed for summative purposes, portfolio assessment can also help in aligning teaching and assessment to facilitate productive learning (Klenowski, 2002). By and large, portfolio content is the key for determining its use orientation, or the purpose for which it is established. Toward this aim, Weigle (2002) attempts to formulate five questions to create a context for writing portfolio assessment: (1) who decides what goes into the portfolio, (2) what types of writing should be included in the portfolio? (3) How many pieces should go into the portfolio? (4) What should be included in the portfolio in addition to the students' classroom writing samples? (5) How can the authorship of the portfolio content be authenticated? (pp. 212-213). As for the first, second, and third questions, the meaning is clear. Concerning the fourth question, this can be expanded by asking whether the portfolio should include only the best work, a range of works from a variety of genres, both in-class and out-of-class works, complete or incomplete piece, and so on. Regarding the fifth question, the authorship of the portfolio content can be authenticated through for example collecting students' writings produced in classroom under teachers' supervision or asking students to provide the development of their written piece i.e. all the versions the students had written to produce the final version.

As the portfolio content is tightly related to the assessment needs, whether for summative or for learning purposes, these questions do not seem to have fixed set answers and vary depending on the purpose of developing the portfolio. According to Arter and Spandel (1992), one of the major rules when dealing with the portfolio is to determine the purpose(s) for which it is developed otherwise the portfolio is just a folder of student work. Most importantly, clarifying the portfolio objectives helps in uncovering ambiguities regarding their implementation (Davies & LeMahieu, 2003). For example, introducing portfolio for learning purposes defines the role the teacher and students should fulfill. Teachers can guide selection, reflection and assessment while students can be the executive agents.

To summarize, portfolios are gatherings of students' work that can document their progress and provide a realistic report on their strengths and weaknesses over time. If organized according to a set of criteria, they could be well developed to not only aid in the creation of a discussion forum between teachers and students but also to highlight the goals to accompany students in their writing. As a result, they can become portfolio for AfL and not only for fragmented and random assessment. In what follows, we discuss how portfolios can serve AfL.

1.8.2 Portfolio for Assessment for Learning

Despite being well-known in the literature, portfolio implementation is not always clear (Gipps, 1994; Tillema, 1998). Smith and Tellima (2003) concur that there are many ambiguities and conflicts over portfolio use despite having interesting and diverse purposes. This might be explained by the large range of objectives portfolios can be used for, since each classroom has distinct goals that need a set of specific criteria to make a given portfolio operates in accordance with them. The main idea about portfolios is that they function clearly to emphasize the many purposes, usages, and views that they may achieve when employed in an educational setting. This might occur as a result of identifying a specific instructional and assessment purpose to direct them toward providing useful feedback.

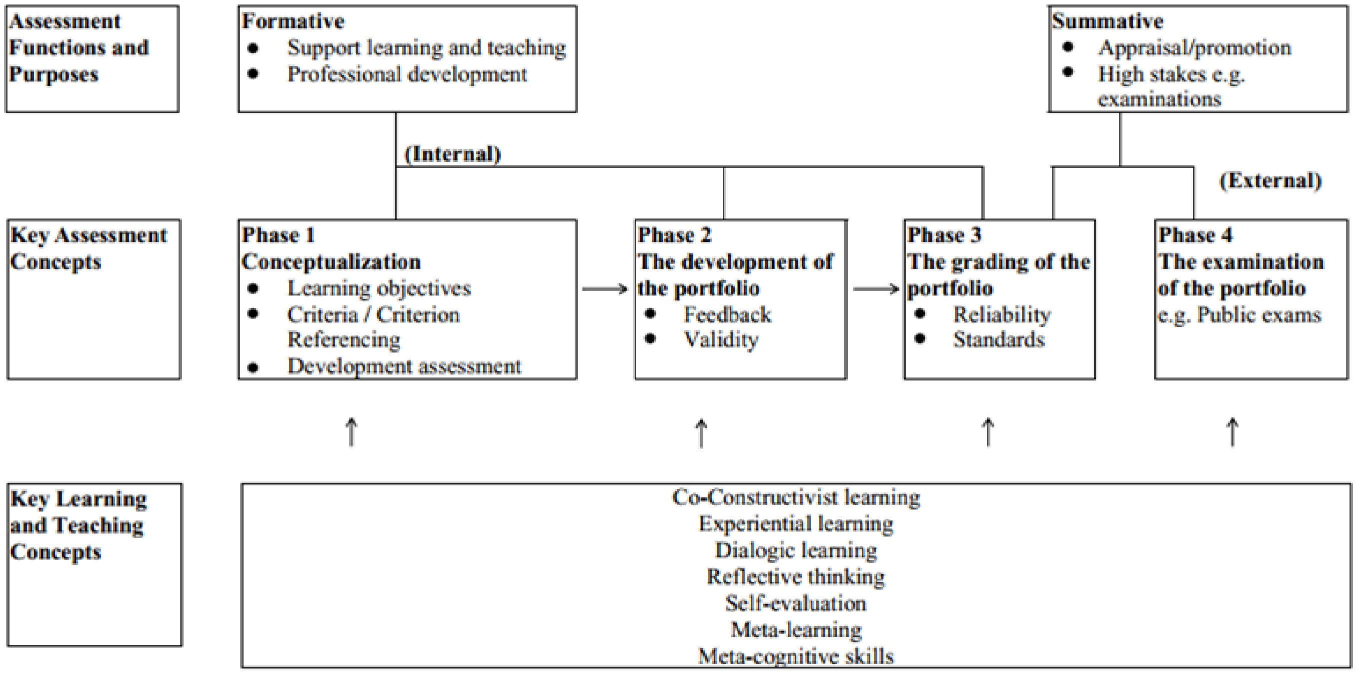
In fact, portfolios were primarily recognized to provide a summative function by accounting for students' writing successes, since their major goal was measuring student achievement because they could record work progress (Genesee & Upshur, 1996). However, there is a growing interest in the use of portfolios for formative purposes as a viable evaluation tool to feedback and offer learners opportunities to reflect on their learning progress. In other words, they can be employed to promote AfL. This, essentially, implies delivering feedback information that encourages rather than merely evaluate classroom learning (Klenowski, 2002). In practice, however, the formative usage of portfolios was noticed anecdotally (Hamp-Lyons & Codon, 2000), which requires further empirical education research to enhance the numerous positions maintained in favor of portfolios' roles (Hamp-Lyons, 1996).

On the other hand, portfolio-based classrooms are "more student-centered, collaborative, and holistic" (Genesee & Upshur, 1996, p. 99) than test-based classrooms. This is to say that portfolios has the potential to foster learning through assessment. Spendlove (2009), for example, believes that portfolios, as part of a holistic strategy, may polish learning and evaluation to offer the learner a substantial profile. Also, they may highly promote learner- centeredness in classroom-based assessment by positioning the student in the center of the classroom rather of being eccentric and ignored. As a result, it is necessary to shift from a teacher-centered approach in a receptive-transmission paradigm, in which the teacher is assumed to be an expert in a certain subject and convey knowledge to a passive recipient, to a constructivist and co-constructivist approach (Klenowski, Askew, & Carnell, 2006). This may be considered a strong evidence for activating the use of portfolio to help in recording students' development, growth, or success. This may endorse portfolios to be used an alternative to fight against traditional and quantitative assessment methods (Smith & Tellima, 2003).

Using a portfolio for learning objectives requires shifting the emphasis from evidence gathering (Arter & Spandel, 1992) to analysis and integration of learning (Klenowski et al., 2006). As a response to the two major limitations identified in writing assessment, writing under timed conditions on unfamiliar topics and assessing a single written sample (Weigle, 2002), portfolio can be viewed as a solution because it does not focus on a single piece of writing as it provides a space for discussing the progress or achievement made over time, exposing the ups and downs of students. This happens if portfolios are well designed and carefully assembled to combine training and assessment (Paulson et al., 1991; Davies & LeMahieu, 2003). They can therefore become a potential instrument for generating compelling proof for learning. This can be possible under these five conditions: (1) portfolios can help in the development and growth of student learning over time (Davies & LeMahieu, 2003); (2) portfolios allow engagement and examination of sustained effort and deeper performance, providing opportunities for collaborative assessment and goal-setting and provide opportunities for collaborative assessment and goal-setting (Genesee & Upshur, 1996), (3) through selection, portfolios can document students' dispositions toward and understandings or misunderstandings about learning (Davies & LeMahieu, 2003), (4) Portfolios provide tangible evidence that can be shared with parents, other educators, and other students (Genesee & Upshur, 1996), and (5) portfolios provide a space of interactivity between students and a chance to reflect upon their own works (Davies & LeMahieu, 2003), through promoting ownership and responsibility for self-reflection to think critically (Paulson et al., 1991; Genesee & Upshur, 1996).

In order to determine creating portfolios' significance that contributes both internally and externally to empowering learning, Klenowski et al. (2006) provided a framework in which they analyze the aims of portfolio use and important evaluation ideas (see Figure 1.2). The authors highlighted the relationship between important concepts such as co-constructivism, formative and SA goals, and dialogic learning in connection to the many stages of portfolio building.

Figure 1.2 *A Framework for Using Portfolios for Learning and Assessment by Klenowski, Askew, and Carnell (2006, p. 268)*


Assessment Functions and Purposes
Formative
Support learning and teaching
Professional development
Summative
Appraisal/promotion
High stakes e.g. examinations
(Internal)
(External)
Key Assessment Concepts
Phase 1
Conceptualization
Learning objectives
Criteria / Criterion Referencing
Development assessment
Phase 2
The development of the portfolio
Feedback
Validity
Phase 3
The grading of the portfolio
Reliability
Standards
Phase 4
The examination of the portfolio
e.g. Public exams
Key Learning and Teaching Concepts
Co-Constructivist learning
Experiential learning
Dialogic learning
Reflective thinking
Self-evaluation
Meta-learning
Meta-cognitive skills

According to this framework, the internal formulation of portfolios helps in appraising the FA purposes, whereas the external outcomes are directly related to the summative ones. Students can advance in language skill development by promoting their critical abilities through various methods of portfolio construction, selection, self-reflection, appropriate criteria, and appropriate decision establishment. Furthermore, feedback and validity are determined to be significant in the development phase when directing portfolio function to be especially for formative purposes. It is worth noting that the authors perceived FA as using assessment to promote and enhance learning rather than employing assessment on a regular basis regardless of its goal. This dilemmatic stance has been fully discussed in Section 1.1 at the beginning of this chapter.

In conclusion, when used to promote learning, portfolios may be used as a means for initiating and developing thoughtful criticisms, which can create an atmosphere where students can be invovled productively. That might work better if the goal is to emphasize the value of feedback above accountability in order to track development along a continuum.Therefore, portfolios can serve as a window through which all members can see how the process works (Paulson et al., 1991). Despite the importance of using portfolios in learning since they focus primarily on the objectives stated, researching different types may provide additional information on how to orient portfolios to cover the many arrays of learning benefits.

1.8.3 Portfolios' Classification

Portfolios are defined in the literature in terms of their purpose or contents. They can also be described in terms of how they are reviewed, such as creating a rubric against which their content will be evaluated or being evaluated without any pre-set criteria. In fact, portfolios are mostly classified according to their function. The most common classifications in the literature are progress/working, best-work, and achievement portfolio. Because our primary focus is on the working portfolio, we will only present the most well-known classification, the one by Lam (2018). However, we may direct the reader to additional classifications, such as those developed by Nitko (1996), Smith and Tellima (2001), and Spendlove (2009).

Lam (2018) identified three kinds of portfolios in term of the purpose, rationale, design and content in the writing context as follows:

Table 1.5 *Classification of Assessment Portfolio. From Lam (2018, p.76)*

	Progress portfolios (growth)	**Working portfolios (efforts)**	**Showcase portfolios (achievements)**
Purpose	Mainly diagnostic; partially formative; minimally summative	Mainly formative; partially summative	Mainly summative; partially formative
Rationale	Keep track of student writing development ;	Celebrate student efforts in writing;	Demonstrate student best writing ability via

	nurture growth in learning writing; promote learner agency	assist students to achieve learning goals and foster ownership in learning	representative work; showcase learning achievements
Design	Longitudinal; sustainable; process-based or product-based	Developmental; Reflective; workshop-like	Autonomous; Metacognitive Emphasize learner choice in writing
Content	Flexible; open-ended; artefacts include pop quizzes, examinations, interim drafts, reflective pieces	Embrace a wide range of learning evidence including unfinished works; Work-in-progress; journal entries	Mainly final products of best entries; Reflective pieces

Despite his efforts to specify each kind, Lam (2018) claims that there are no firm prescriptions for defining the features of a particular portfolio task type. The inclusion of portfolio tasks depends particularly on the objective of a portfolio assessment program, essentially when the assessment outcomes are intended to guide the teaching and learning of writing without constraints. According to Weigle (2002), whatever the portfolio is, it can include writing drafts or final products, and regardless how it is collected: strict guidelines or left to students' discretion, it cannot be called such unless it contains more than one written piece because the goal is to provide evidence on students' writing performance, and one piece is insufficient.

Regardless the different types and the various purposes portfolios may fulfill and achieve, the important rule is to determine which type or types to exploit in a given classroom. In our study, we are concerned with a manual working portfolio that allows students to reflect on their work and backtrack their development in order to identify their strengths and flaws.

1.8.4 Working Portfolio

The term ‘working portfolio’ varies from one researcher to another since everyone categorizes portfolios and defines them based on his/her own conception. To avoid being limited by such labels, in this study, we are concerned with the portfolio that monitors students’ development, which can be referred to as working, growth, effort, or learning. For preference sake, we used the phrase ‘working portfolio’. To consider the unturned appellations of different portfolios’ types, one can consider Lam (2018)’s categorization. The author evoked growth portfolio and working portfolio as two distinct types, whereas authors such as Danielson and Abrutyn (1997), Campbell, Cignetti, Melenyzer, Nettles, and Wyman (1997), and Rolheiser, Bower, and Stevahn (2000) used the word working to refer to the type that tracks students’ learning and progress. As a result, we see that discussing those appellations is pointless, and we must instead focus on the purpose of the portfolio we want to use to meet our goal.

The working portfolio is defined as an intentional collection of student works driven by learning objectives (Danielson & Abrutyn, 1997) that is used to track students’ efforts (Lam, 2018). The working portfolio, also known as the learning portfolio (Burner, 2014), is primarily used for FA but may also be used for grading and SA ends (Lam, 2018). As a result, it may be used to assess students’ requirements (Danielson & Abrutyn, 1997), orientations, preferences, and responsiveness. This is why there seems to be agreement on the idea that the collection in the working portfolio is made up of all of the students’ works that are related to a pre-determined goal (Danielson & Abrutyn, 1997). The works included in the portfolio may be finished or unfinished. The goal of the reflection phase is to monitor and examine the whole portfolio development to identify the strengths and flaws of student works. As a result, employing pre-set criteria to simplify the work might be really beneficial. This implementation can be scaffolded by either teachers and/or peers.

1.8.5 Researching Portfolios

The field in which portfolios are introduced is crucial since a lack of research in that area might have a significant impact on its proper implementation. For example, studies on how to use portfolios for learning to assist students develop and understand their own learning, assessment, and future professional practices are scarce in higher education (Klenowski et al., 2006). To alert teachers or researchers to some of the issues that may emerge during portfolio implementation, Cappeli proposed four warnings: (1) there is a need for scoring guides that reflect the skills and abilities we want to measure, (2) portfolios may not capture a broad enough variety of skills, (3) there is a lack of reliability and validity, and (4) logistical problems exist. Despite not largely addressing summative purposes, Cappeli's (As cited in Hamp-Lyons & Condon, 2000, p. 19-20) cautions should be acknowledged and taken into account before approaching this field since they help in optimizing efforts toward striving for more successful implementation.

Researching portfolios is a topic that encounters various difficulties and ambiguities due to extended discussion, or lack thereof, regarding their design, content, rationale, and aims. Zeichner and Wray (2001) posit that individuals engaging portfolios require more understanding regarding the many ways in which they (portfolios) have been employed. The focus here is not primarily examining such conflicts and confusions, since this is a distinct axis of research, but one must bear in mind that there is no ideal way for constructing, designing, evaluating, and exploiting portfolios. For practical purposes while feeling blocked when approaching portfolios, a good rule of thumb to follow is asking the following question: Why am I using this portfolio? This question is inextricably linked to the goal of developing a given portfolio. Therefore, addressing this question may provide solutions and further instructions, paving the path for more insightful and effective implementation.

Regardless of the problems and ambiguities associated with researching portfolios, scholars (e.g., Hamp-Lyons & Condon, 2000; Genesee & Upshur, 1996) have provided a set of criteria for implementing portfolios in classroom training. Despite the fact that the number of phases described varies by source, all frameworks have common proposals summarized in: (1) defining the aim, (2) developing an overall design and structure, (3) selecting the content, and (3) deciding on and constructing an assessment strategy.

Upgrading portfolios to be interactive vehicles for promoting student involvement in learning, Genesee and Upshur (1996) suggested a number of further useful guidelines: (a) involving students in making decision about portfolios whenever possible, (b) scheduling portfolio conferences periodically to give students opportunities to discuss their works with their teachers and try to formulate individual goals, (c) fostering communication between students where they share their works and discuss them via providing constructive feedback, and (d) always adopting a positive, collaborative, and supportive attitude (Genesee & Upshur, 1996). The first guideline may cover how, why, when, and what to analyze, as well as which criteria and why to choose which works to include and how to assemble and create the portfolio. Sharing it with students who have previously used portfolios as well as those who have never used them is beneficial to both. For the former, doing so may help them reestablish and review what they have learned with portfolios and understand how to use them in the current situation, or may expand their learning if the principles followed are comparable to those they followed previously. However, for the former students can require instruction and careful assistance to learn how to implement portfolio.

In a nutshell, investigating portfolios is an open topic for more contributions since many aspects, including implementation process, goal, content, and rationale, have not yet been completely examined. Researching portfolios overall as a field of research is tightly related to struggling with portfolio construction, as this basically direct the tool toward the implementation procedures and aims to achieve. This essentially creates ambiguities for data collectors and tenses the inherent relationship between evidence selection and reflection.

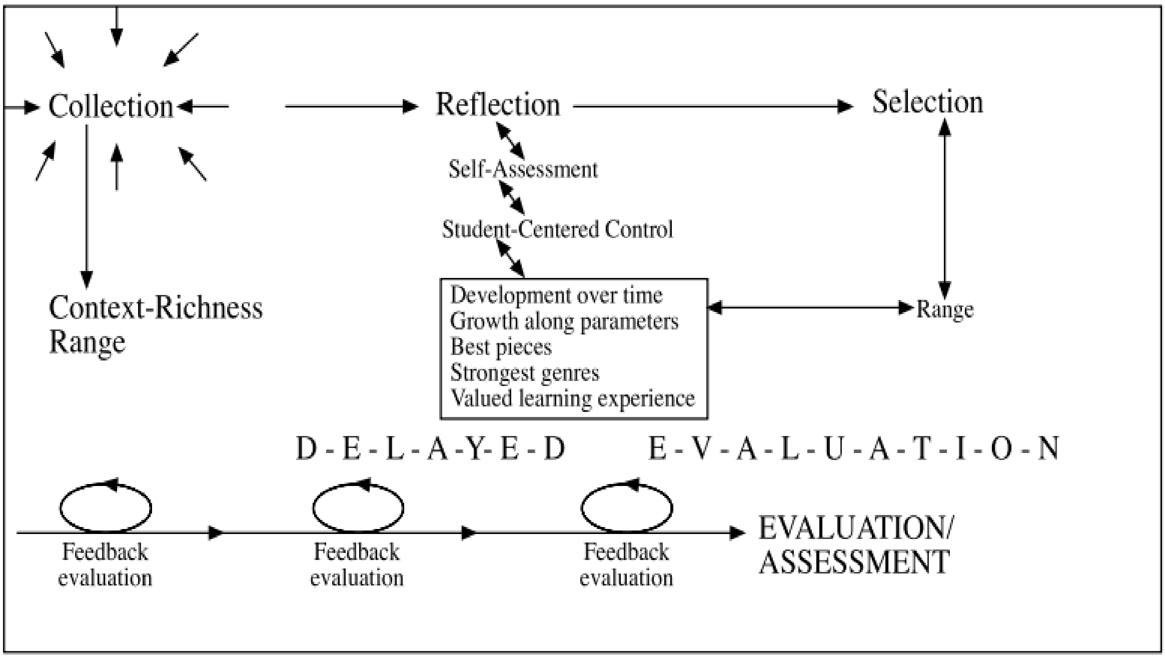
1.8.6 Portfolio Assessment in Writing

Writing portfolio assessment has lately been promoted because of its formative potential, which can boost learning through repeated opportunities to edit drafts and give rich feedback assistance, via self, peer, and teacher feedback. It is believed that portfolios can achieve an objective that many other assessment systems cannot since they shift the student’s position in evaluation from passive research participant to active participant (Lam, 2018). In a typical portfolio development process, students are expected to collect, select and reflect upon works in progress under the guidance of teachers or other communal resources (Burner, 2014).

It is important to consider that portfolio assessment has seen numerous means of collection, assessment, and usage since it is extensively employed in diverse situations; nonetheless, there are several traits that are found common in “many, if not most, portfolio assessment programs” (Weigle, 2002, p.198). In line with this, Hamp-Lyons and Condon (2000) produced a seminal study on portfolio construction and assessment that is still used as a reference. They see that portfolios include nine qualities: collection, range, context richness, delayed evaluation, selection, student centered control, reflection and self-assessment, growth along defined parameters, and development through. Among these nine portfolio traits, following the two authors collection, selection, and reflection are the most significant.

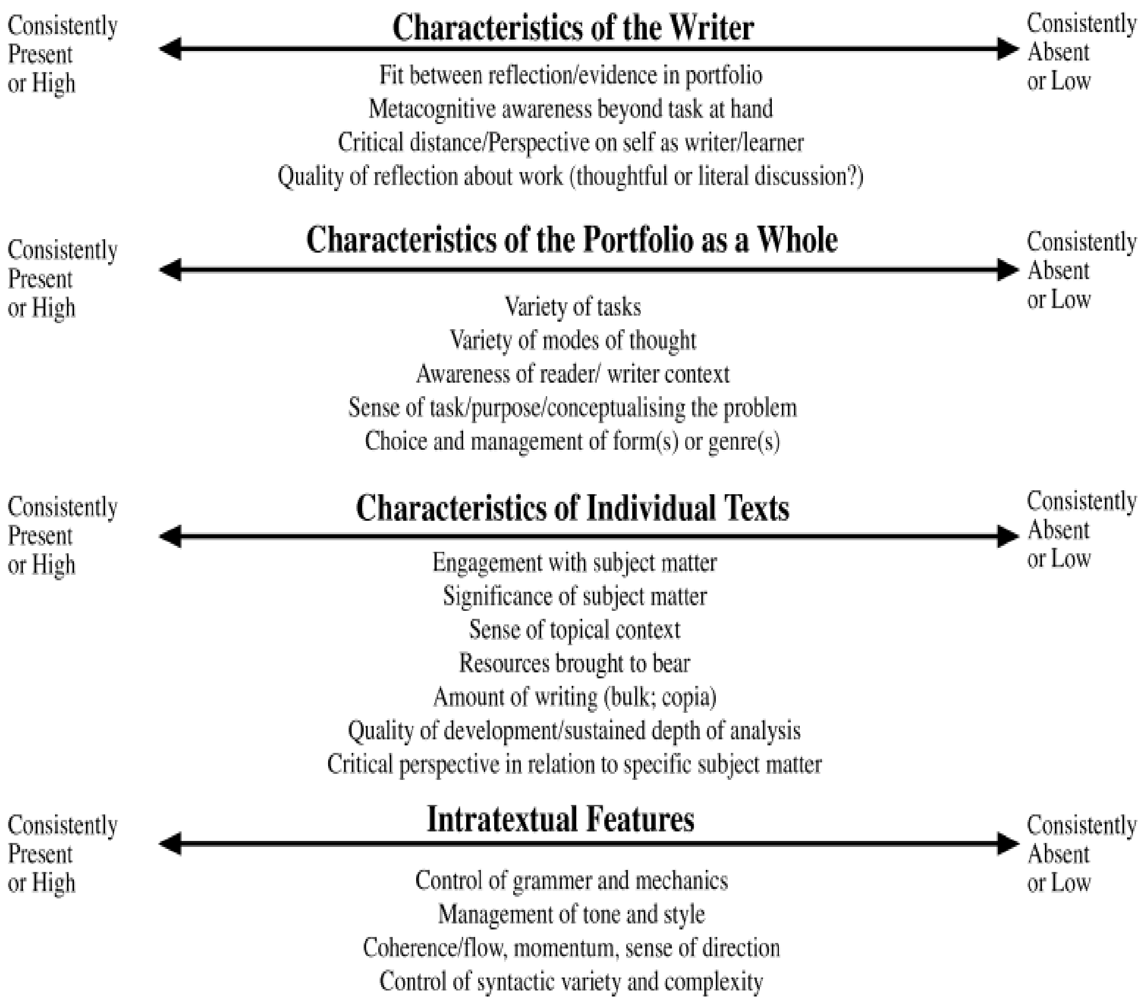
For collection, the portfolio should not include only one written piece; otherwise, the portfolio cannot be called as such. It might include final product or early drafts and can be gathered according to specific criteria or students’ discretion. Some helpful tips are proposed to make the collecting process goes as easily as possible. Genesee and Upshur (1996), for example, believe that the number of works included in a portfolio should be limited for practical reasons such as being easily stored, reviewed, and assessed. As collection alone is not enough to make a portfolio work, reflection and selection are essential as well (Weigle, 2002). According to Hamp-Lyons and Condon (2000), without reflection, the portfolio cannot function as a pedagogical tool, teacher development tool, or evaluation tool, but is just a pile or a large folder. Without the delayed evaluation component in which students are given a chance to reflect, revise, and select their writing, students will be less motivated to work on their portfolio and will regard it as a pointless work. Combining the importance of the three basic features of a portfolio, Weigle (2002) argues that intentional selection and reflection as well as the organization of the specific content is what turns the collection of written samples into a portfolio. These characteristics are best illustrated in Figure 1.3.

Figure 1.3 *Basic Portfolio Characteristics. From Hamp-Lyons and Condon (2000, p.122)*


Collection
Reflection
Selection
Self-Assessment
Student-Centered Control
Context-Richness
Range
Development over time
Growth along parameters
Best pieces
Strongest genres
Valued learning experience
Range
D-E-L-A-Y-E-D E-V-A-L-U-A-T-I-O-N
Feedback evaluation
Feedback evaluation
Feedback evaluation
EVALUATION/ ASSESSMENT

To assess writing portfolio, Hamp-Lyons and Codon (2000), proposed the diagram in Figure 1.4 which includes the main axes that can be taken into account in the assessment process. They proposed four axes in the project of Denver and Michigan universities for first year composition teachers for the aim of developing the assessment criteria, and they argued that the best portfolio works as a whole rather than in terms of the parts included in.

Figure 1.4 *Dimensions for Assessing the Portfolio. From Hamp-Lyons and Codon (2000, p. 144)*


Consistently Present or High
Characteristics of the Writer
Consistently Absent or Low
Fit between reflection/evidence in portfolio
Metacognitive awareness beyond task at hand
Critical distance/Perspective on self as writer/learner
Quality of reflection about work (thoughtful or literal discussion?)
Consistently Present or High
Characteristics of the Portfolio as a Whole
Consistently Absent or Low
Variety of tasks
Variety of modes of thought
Awareness of reader/ writer context
Sense of task/purpose/conceptualising the problem
Choice and management of form(s) or genre(s)
Consistently Present or High
Characteristics of Individual Texts
Consistently Absent or Low
Engagement with subject matter
Significance of subject matter
Sense of topical context
Resources brought to bear
Amount of writing (bulk; copia)
Quality of development/sustained depth of analysis
Critical perspective in relation to specific subject matter
Consistently Present or High
Intratextual Features
Consistently Absent or Low
Control of grammer and mechanics
Management of tone and style
Coherence/flow, momentum, sense of direction
Control of syntactic variety and complexity

In light of Figure 1.4, the axes included in assessing portfolios are four in number: characteristics of the writer, the portfolio as a whole, individual texts, and intertextual features. Each axis includes a set of characteristics which describe and define some qualities that should be included, and these characteristics range from low to high in terms of their absence or presence. The authors positioned the writers' function, as indicated on the axis of writer's characteristics, in assessment and identified a set of attributes they must retain for subsequent assessment such as metacognitive awareness. In the axis about the characteristics of the portfolio as a whole, the authors highlighted the main points to consider when constructing the portfolio. In the third axis about the characteristics of individual texts, the authors described the texts that should be included in the whole portfolio, and in the fourth axis about intertextual features, the authors listed a set of assessment criteria.

It is worth bearing in mind that this project took more than ten years to come into existence, and the assessment aimed at behind was summative. As we are primarily interested in portfolios for learning purposes rather than pure grading, this model will not be relied upon entirely, but it is very insightful and helpful, particularly when considering aspects that could directly work for AfL purposes, such as the importance of involving the writer in the assessment process. Some criticism may be addressed to this model such as positioning the student writer. This latter can be involved in all phases of constructing the portfolio and not only in reflection/evidence phase. As we have discussed so far, students have a major role to play in AfL especially when giving feedback and developing the assessment criteria.

After reviewing the literature on writing portfolios, the conclusion reached is that the primary focus of employing portfolios recently is entirely for learning purposes. This is done in the hope of increasing students’ awareness of their writing progress and motivating them to work more and better. Despite the inconclusive results regarding the guidelines to follow while constructing a portfolio, this tool remains an opportunity for students to become active participants. A rule of thumb to follow before deciding upon using a portfolio is underlining what purpose is used for to select the appropriate type, content, and reflection strategy.

Part III. Writing in L2 Context

This section examines the literature related to the study’s dependent variable, EFL writing. At the outset, we clarify that L1 and L2 are symbols used to refer to the native setting and ESL and/or EFL context, respectively. A closer look at these two symbols, L1 and L2, reveals that researchers did not agree on a given decision like L2 is strictly ESL, EFL, or both because many used these two terms interchangeably and referred to them as L2 while others differentiated between them. In this context, Matsuda, Ortmeier-Hooper, and Matsuda (2009) claim that the distinction between first, second, and third languages “has become problematic” (Matsuda, Ortmeier-Hooper, &

Matsuda, 2009, p. 458). In fact, we are not interested in delving into this topic because it is a subject of multilingualism and this latter is a distinct area of research. However, we just alert the reader that we are employing L2 or EFL in this study to insist that the context of this study is non-native.

The review in this section is divided into two main parts: the first one is devoted to approaching L2 writing by defining writing, examining briefly the difference between Arabic and English writing, researching writing, relationship between L1 and L2 writing, approaches to teaching L1 and L2 writing, and difficulties and challenges faced by Algerian EFL students when writing in English while the second part is devoted to paragraph writing and the writing criteria required to decide that a given paragraph adheres to the standards..

1.9 Approaching L2 Writing

1.9.1 Writing Defined

Writing is difficult to define due to the various meanings this word has in different English dictionaries because of its history, purpose, teaching and learning philosophies and pedagogies, and following the various perceptions that everyone has about it, whether it is isolated, connected, or embedded within a set of dimensions (s). In order to understand the meaning of this notion, we have tried to discuss a collection of definitions. The first definition worth presenting is the one by Coulmas (2003). He postulated six meanings of writing, beginning with seeing writing according to its form and moving to its function:

- a system of recording language by means of visible or tactile marks
- the activity of putting such a system to use
- the result of such an activity, a text
- the particular form of such a result , a script style such as block letter writing
- artistic composition. This idea is also supported by Hedge (1988) and White (1987) who see that writing is a discovery and a creative process.
- a professional occupation (p.1).

Writing is viewed as a form of communication to convey thoughts and feelings (Harmer, 2004) since it is a thinking process (Sinclair, 2010) and the consequence of a collection of cognitive operations (Hedge, 1988). In the EFL context, Shokrpour and Fallahzadeh (2007) characterize writing as "a complex activity, a social act which reflects the writer's communicative skills which is difficult to develop and learn" (p. 147). This is to say that writing that serves a communicative purpose is a difficult task because it is not only a complex activity that requires sophisticated cognitive abilities, but it is also a socially grounded construct that needs an intricate set of factors, including elements of society, such as culture, to achieve its purpose.

Rather than defining it conventionally, some scholars believe that writing includes components that must be owned in order to stand out. According to Raimes (1983), writing should encompass content, the writer's method, audience, goal, word choice, mechanics, grammar, and context. These criteria are reflected in figure 1.4 in Section 1.8.6. However, for Grabe and Kaplan (1996) writing rhetorically includes the reader, the recipient of the final product, the writer, and originator of the message; the subject matter and the text itself. Finally, Myhill (2009) sees that L1 writing includes three main perspectives: a cognitive psychological perspective, a socio-cultural perspective, and a linguistic perspective (p.405).

Based on the above descriptions, writing appears to be a complicated task rather than an isolated practice, requiring a significant amount of time and effort from both parts, teachers and students, to understand and master its components. In addition, multiple components should be incorporated to make it a reflective activity (Chakraverty & Gautum, 2000) and a complex process. Not only does it require "imagination, feeling, state of mood, cognitive state, capability with the medium, context and other factors" (Andrews, 2001, p. 43), writing needs also further competencies of analysis and classification of any background knowledge (Chakraverty & Gautum, 2000).

1.9.2 Difference between Arabic and English Writing

Arabic and English are different because they come from two different families and different social, cultural and linguistic backgrounds. Some differences between Arabic and English languages are figured out as follows:

> Arabic is written from right to left, whereas, English is written from left to right. The actual forming of the letters also presents difficulties for Arab learners. In Arabic, letters are formed by a series of strokes, unlike the continuous flow of the Roman alphabet used for the English script. There is also difference between the printing and script of Arabic. Moreover, Arabic writing conventions and spelling systems are vastly different from those of English. For example there are no capitals in Arabic and prepositions are joined the word that follows them. Arabic is also a very phonetic language and the variations in vowel sounds found in English (for example, the sound of the vowel ‘a’ is car, make and bat) and in diphthongs cause difficulties for the Arab learner’. (Watson, 2004, p. 42-43)

It was found that when two languages are different in a statement, the old habit interferes with the new structure, causing the learner to struggle with transfer; however, when the message is the same, the first language may help in the acquisition of the new language (Ellis, 2015). So we may assume that the interference between French and English is less than the one between Arabic and English or French because the statement is very different. For this reason, Abi Samra (2003) found that nearly one third of the Arab L2 learners’ errors can be caused by transfer from L1 and over-application of L2.

1.9.3 Researching L2 Writing

The interest in L2 writing began in the 1960s but did not receive much attention as this latter was overshadowed by the Audio-lingual Approach (Matsuda, 2003). According to Ferris and Hedgcock (2014), it is only until the late of the 1980s that L2 writing became a distinct field of research and it is still lacking “a tidy corpus of conclusive theory and research on which to base a straightforward introduction to processes of learning and teaching.” (p.3). On the other hand, Cumming and Riazi (2000) believe that the field of L2 writing still needs further understanding as people still do not know how to learn and write and teachers lack enough knowledge about how to better teach for better learning in L2 writing.

Despite being extensively researched, no models for L2 writing have been found in the literature, and researchers in this area suggest applying L1 models with suitable adjustments (Hyland, 2003). Similarly, while some writers, such as Grabe and Kaplan (1996), distinguish between L1 and L2, others believe that these distinctions are exaggerated (Friedlander, 1990). For further clarification, Matsumoto (1995) suggested that there are fundamental and common concerns to every act of writing regardless the language to use or produce, and that less proficient writer should not be hindered by the differences. Despite this, L2 writing cannot exist as a distinct discipline because of the ongoing advancement in the L1 field. This implies that researchers cannot arrive at a distinct theory or model since they are confronted with another theory or model that has been freshly created or updated in L1 research.

To simplify matters and give an entry point into this subject, a number of researchers, such as Raimes (1993) and Hyland (2009), suggest that teaching and researching writing may be undertaken according to three basic approaches: text-oriented, writer-oriented, and reader-oriented. These approaches have an influence on teaching and researching writing in L2 contexts since they are well-known in L1 writing situations. Silva (1990), on the other hand, suggests four components for building L2 writing models: 1) L2 writing theory; 2) L2 writing nature research; 3) L2 writing instruction theory research; and 4) L2 writing practice. As we can see, approaching L2 writing cannot be fully linked to the L1 field since the L2 context has unique characteristics. For example, practicing L2 writing requires an L2 context and this latter is completely interwoven with dimensions like culture and society. To distinguish L2 writing, the difference between L1 and L2 is viewed important. For this purpose, Cohen and Cavalcanti (1990) proposed a strategy based on ethnographic data and the context of each language. This is applied to L2 language as well. After discussing the problems of investigating L2 writing, we will look at how L1 and L2 writing might be connected.

1.9.4 Relationship between L1 and L2 Writing

It has been concluded that L2 writers approaches writing differently from their L1 counterparts (Silva, 1990; Weigle, 2002; Hyland, 2009), but the necessity of L2 writers to the L1 ones is critical in order to regulate L2 writers' mistranslations of L1 structure and terminology (Silva, 1990). This explains the vast range of issues that L2 writers have, and the due attention researchers should put on those problems to solve in order to run L2 writing smoothly and much easily. Form this perspective, the relationship between the two is raised to show how much contribution L1 writing can offer for L2 one.

To examine the problems figured out in L2 writing, Matsuda (2003), for example, related this to the fact that L2 writers plan less, employ less goal-setting, and review less, whereas Grabe and Kaplan (1996) contend that L2 writers are irritated by micro-errors, grammar and language form, at the cost of macro-errors, content. They, consequently, find difficulties in making their text appropriate in length, especially short, and fluent with fewer errors (Matsuda et al., 2009). Indeed, the problems in L2 writing may be linked to the complexity generated by making all of the different components function together, and creating a text is not the result of a set of separate stages rather than of interconnected stages within a non-mechanistic process. One of the problems generated from dealing with this process for L2 writers is producing long texts. Regarding this, Grabe and Kaplan's (1996) position, for example, is subject to discussion. The authors see that is caused by repeated use of coordination and the preference of this latter over subordination. This can also be explained by the limited vocabulary of L2 writers and the difficulties of literal translation from L1 to L2. Limited vocabulary knowledge may hinder L2 writers' ability to generate written works with rich language and context-related meaning, and translating from L1 to L2 with a limited vocabulary may result in writings that lack accuracy and fluency. In the context of translation challenge, failing to consider the structural differences between L1 and L2 languages, considering Arabic redundant and indirect in comparison to English, is basically among the main reason of producing lengthy texts.

The debate on this topic may be too extensive; therefore, we shall limit the discussion to some conflicting points in teaching L2 writing from a multilingual perspective. Some scholars, (e.g., Canagarajah & Jerskey, 2009), believe that transitioning from L1 to L2 writing improves the quality of L2 texts, but others believe that the writers' first language interferes with L2 writing (e.g. Wang, 2003; Derakhshan & Karimi, 2015). On the other hand, some scholars suggest that writers who did not develop strong methods and strategies in their native language would be unable to transfer any into L2 writing (e.g., Kroll, 1990; Grabe & Kaplan, 1996), as L2 writers may retrieve information about L2 themes from their first language background (Silva, 1990). In his study, Yanqun (2009) concluded that using L1 had no effect on L2 writing; therefore, he urged teachers not to discourage their students from thinking in their first language. Basically, the major aims of L2 authors is to improve their content and structure through translation (Cumming, 1989); nevertheless, teaching L2 should not be based on word-for-word translation more than being “larger chunks gradually getting beyond the words” (Yanqun, 2009, p.13). Another thing to consider is writing on the chosen topic. This would assist L2 authors in avoiding what Smith (2013) refers to as ‘writing block’. This occurs when L2 writers forget their ideas, do not write them down because they are scared of judging or being judged, or do not want to disappoint their imaginary readers.

Finally, we can argue that the link between L1 and L2 is certainly obvious yet complex, and this is one of the key challenges in teaching L2 writing. This problem is exacerbated by the many levels seen in L2 writing classes. For this reason, various approaches to teaching writings have emerged.

1.9.5 Approaches to Teaching L1 and L2 Writing

Teaching writing is very contentious since there is no single and unique strategy. It may focus on the purpose, which will depend on accuracy, fluency, and both, target the writing process itself, which focuses on the stage through which the writing goes through, or be a combination of the two. In what follows, we present and discuss briefly the main approaches used in teaching writing.

1.9.5.1 Controlled and Free Composition. Some researchers supported Controlled Composition, which evolved from Skinner’s Audio-lingual method in which short and competent written pieces were used to practice specific syntactic and lexical patterns (Silva, 1990; Kroll, 2001), over Free Composition, which represents students’ own texts with errors and expects grammar to develop over time. In contrast to the Free Composition, the Controlled method prioritizes accuracy above fluency. It encourages practicing exercises, descriptive analysis, and error analysis by teachers. For further understanding, Crookes and Chaudron (1991) attempted to outline the distinctions between the two in Table 1.6.

Table 1.6 *Difference between Controlled and Free Composition by Crookes and Chaudron (1991, p. 52)*

Controlled Composition	Free Composition
Teacher-centered	Student-centered
Manipulative	Communicative
Structured	Open-ended
Predicted student response	Unpredicted responses
Pre-planned objectives	Negotiated objectives
Set curriculum	Cooperative curriculum

It is worth mentioning that both techniques appear under Product-based Writing Approach. For information, Controlled Composition and Current-Traditional Methods both emphasize discursive form (Silva, 1990; Matsuda, 1999), traditional form (Silva, 1990; Kroll, 2001), and Current-Traditional Rhetoric (Silva, 1990; Kroll, 2001). (Silva, 1990; Grabe & Kaplan, 1996). These concepts are explained in the forthcoming sections.

1.9.5.2 Current-Traditional Approach. The Current-traditional Approach is referred to as Paragraph-pattern Approach by Raimes (1983) and Current-traditional Rhetoric by Grabe and Kaplan (1996). It was introduced in second and foreign language study to meet the requirement for long written conversation (Silva, 1990). Under this approach, students were taught to compose coherent pieces by arranging and combining sentences into paragraphs according to a predetermined formula. It fundamentally emphasizes combining the syntactic rules with the content while accounting for variations between L1 and L2 in what is called Contrastive Rhetoric Method (Grabe & Kaplan, 1996).

Overall, this method is guiding composition focusing on “topic sentence, supporting sentences, concluding sentences, and transitions”, and covering “essay development” and “paragraph principles to larger stretches of discourse” (Silva, 1990, p.14). Students’ focus is on composing product rather than process. Instead of generating individual sentences, the Current-Traditional Approach was developed to arrange the content by constructing paragraphs and then essays. This is a product-oriented strategy. Because the Current-Traditional Approach concentrates on the form of language and cannot adequately assist students in L2 writing, the Process Approach emerged in the 1980s to dominate writing theories (Grabe & Kaplan, 1996).

1.9.5.3 Product Approach. This approach is viewed by Tribble (1996) as a basic linear model of the writing process, progressing systematically from prewriting to composing and subsequently to correcting, but Raimes (1985) contends that writing is recursive. It is used to reinforce L2 writing in terms of grammatical and syntactical forms. It emphasizes employing model paragraphs, sentence exercises, and other grammar and syntax-related activities in general. This approach was criticized by authors such as Zamel (1982), Raimes (1987), and Prodromou (1995). It was criticized for placing too much emphasis on grammar and syntax while ignoring the aim of writing and the audience, as well as a lack of motivation on the part of the learners.

The product approach is typically concerned with the final written product rather than how it is generated. According to Prodromou (1995), the learners’ position is not valued, which impacts their both personal and linguistic potentials, in addition to the permanent corrections that influence students’ motivation and self-esteem. This approach was not completely rejected because at some time there is a final product which needs further linguistic refinement. However, the interest shifted from what to produce to how to produce through the Writing Process Approach, also called Process-oriented Writing

1.9.5.4 Writing Process and Process Writing. To keep the upcoming discussion on track, a perplexing issue should be addressed and answered. What, if any, distinction exists between writing process and process writing? The writing process is seen as consisting of four major stages: planning, drafting, revising, and editing, whereas process writing is a term used in L2 classrooms to mean teaching writing through the writing process approach without distancing writing from the written product (Seow, 2002) “to construct process-oriented writing instruction that will affect performance” (Freedman, Dyson, Flower, & Chafe, 1987, p. 13). In addition to the four core stages of the writing process, this approach incorporates “additional phases externally forced on students by the instructor, namely, reacting (sharing), assessing, and post-writing (Seow, 2002). In brief, process writing focuses on teaching writing while avoiding any linearity in the process and connecting the writing process to the end written result.

In fact, the writing process approach emerged as a result of general dissatisfaction with the product approach (Silva, 1990). It is viewed as a non-linear exploratory and creative process in which authors uncover and reformulate their ideas as they strive to approach meaning (Zamel, 1983a). As writing is a complex task that requires knowledge about the text, including form and structure, content, and accurate language, and related aspects, namely audience, purpose, critical thinking and creative potential to compose, writing process exemplifies the concept of writing as a process of generating both order and meaning (Matsuda, 2003). Historically, practice and research have examined the writing process from four perspectives: expressive approach, cognitive process, and genre approach.

a.Expressive Approach. This approach is related to the free writing school founded by Peter Elbow (1973) in ‘Writing without Teachers’. In this writing context, students were given the chance to rely on themselves through free writing (Elbow, 1990a; Faigley, 1986) without limits, i.e. expressing what they thought in whichever way they wanted in order to be creative. This approach was advocated by the L1 research field, but also became relevant to L2 writing instruction because it promotes fluency at the expense of accuracy and puts the writer the primary focus. This strategy may be related to another approach known as the Cognitive Approach. This approach, which first appeared in the 1970s, seeks to connect writing to cognitive psychology and psycholinguistics. Researchers in this field believe that there is interactivity between the writer and the reader, that it is a goal-directed activity, and that skilled writers compose differently than unskilled ones (Bereiter & Scardamalia, 2013). In brief, this approach is a shifted blend of process and product (Reid, 1993).

b. Contrastive Rhetoric Approach. This approach was widely recognized in the 1980s and 1990s. It used discourse and text linguistic research to study the text at the discourse level in order to "understand the varying patterns of organizational preferences in student writing work" (Grabe & Kaplan, 1996, p.28). It is given due attention because it represents a change in teaching writing; going from the sentence level to the discourse level (Kaplan, 1966) would help in understanding what cultural, social, and linguistic patterns can impact L2 writing. When writers from one culture write to readers from another culture, they frequently struggle to grasp the written discourse because they do not share the same perspectives, resulting in "audience expectations" issues and disparities in "cognitive style" rather than "cognitive competence" (Reid, 1993, p. 62). This demonstrates that writing is more than only composing using one's own qualities, including creativity and imagination, but adding the social component in writings has a firm position. Therefore, a genre-based strategy has been proposed. According to Howeritz (1986b), academic discourse on genres plays an important function in immersing learners in the academic milieu and facilitating socialization.

c. Genre Approach. It is based on providing students with the opportunity to explicitly learn about the different genres, particularly their grammatical and discourse characteristics, similarly to "systemic functional linguistics that is concerned with the relationship between language and its social settings" (Hyon, 1996, p. 696). In this context, Canagarajah and Jerskey (2009) observe that "the discourse of a particular genre or disciplinary writing comes with its own grammar and vocabulary and its own community reader/writer" (p. 480). This approach assists L2 students in using the language in real-world communication by defining patterns for a variety of objectives such as learning how to describe and what features characterize this genre. It makes the task easier by breaking down the entire into its constituent parts. The genre-based approach is theoretically underpinned by Vygotsky's (1978) notion of the collaborative nature of learning between the teacher and students; the teacher plays an instrumental role in supporting students to progress from modeling target genres to independently constructing their own texts through drafting and peer assistance. On the other hand, Badger and White (2000) outline the substance of the genre-based writing by emphasizing on linguistic knowledge and social aims.

To feed out the discussion, we see that defining the term genre would offer further understanding to this approach. Toward this aim, two perspectives were offered. The conventional stance regards genre as a collection of distinguishable qualities of a specific work, but the recent one defines genre as “typified rhetorical activities based on recurring occurrences” (Miller, 1984, p. 159). The first definition emphasizes the written form, whereas the second emphasizes the social dimension. Indeed, we have noticed that the genre in writing is broadly spread across the curriculum, syllabus, and disciplines.

In the context of genre study, Hyland (2002) defined three approaches. First, the Sydney Genre School (SGS), whose scholars viewed teaching writing as a process that occurs through stages within a social context to achieve specific communicative textual goals (Martin, 1984), was founded on Systemic Functional Linguistics (Halliday, 1978b; Halliday & Hasan, 1989) which “highlights the relationship between language, text, and context” (Coffin & Donohue, 2012, p. 65). This school provided primary, intermediate, and university education and spotted underprivileged students (Rose & Martin, 2012). The assumption underpinning the second teaching writing approach, or English for Specific/Special Purposes (ESP), is that teaching genre has no explicit reference to any theory of language. Toward this aim, Swales (1990) describes genre’s schematic structure as “a class of communicative events, the members of which share some set of communicative purposes” (p. 58), which are recognized by insiders of a certain discourse community. Both SGS and ESP approaches teach genre explicitly; scaffolding in SGS and a focus on rhetorical moves in ESP enable students comprehend “how target texts are constructed and why they are written the way they are” (Hyland, 2004a, p. 11).

As a third option in a New Rhetorical Approach (NRA), genre is taught “in the process of performing a rhetorical action in its target context of use – which is the situation in disciplinary classroom" (Russell, Lea, Parker, Street, & Donahue, 2009, p. 410) because genre is viewed as a socially standard strategy, integrated within a typical form of discourse to respond to a given rhetorical situation (Coe & Freedman, 1998). Helping students comprehend and critically engage in diverse discourses is a source of strength in genre pedagogies in L2 writing. This is accomplished by taking into account students' requirements and enhancing their resources via analyzing and deconstructing while studying language and context (Hyland, 2007). Despite not seeking to develop models of the writing process and frequently being considered as product rather than process focused (Hyland, 2002), genre studies play an essential role in the area of writing study. As a result, the genre approach provides students with the ability to investigate, compare, and deconstruct cultural and linguistic materials in order to expose their essential assumptions (Hasan, 1996).

This approach is not exempt from criticism, as any approach. For example, Canagarajah and Jersky (2009) investigate how much we generalize to come up with powerful descriptions and definitions for genres in academic contexts as well as the capacity to generalize all texts inside a stereotyped form or if being looked comparable. The rising question is whether it is better to teach grammar and syntax or the genre first. For Ferris (2002), teaching grammar and syntax comes first and then genre, but Knapp and Watkins (1994) believe that everything should go hand in hand without being separated. Discussing this topic is not limited at this level because we can inquire whether including discourse in-between; perceiving a discourse first in the linguistic dimension, or in the social context.

1.9.5.5 Post-Process Approach. Since 2000, L2 writing concerns have shifted to Post Process inquiry, particularly when social, cultural, and educational factors have had a substantial impact on L2 language writing studies (Fujeda, 2006). Indeed, the Post Process Method to teaching writing did not replace the Process Approach, but rather extended it (McComiskey, 2000). It is not a coherent theoretical front, but it may be characterized as the rejection of "process supremacy at the expense of other elements of writing and writing teaching" (Matsuda, 2003). It approaches writing as a whole rather than looking for the processes through which a written work is produced.

However, according to Badger and White (2000), the post-process approach is based on a synthesis of three approaches: product, process, and genre because they are all seen as complementary; writing requires knowledge about language (as in product and genre approaches), knowledge of the writing context and purpose (as in genre approaches), language use skills (as in process approaches), exploitation of students' potential (as in process approaches), and providing input to which students respond (as in product and genre approaches).

To conclude the discussion about approaches to teaching L1 and L2 writing, it seems worth pointing that we have purposefully offered a thorough analysis of the approaches to teaching writing in order to demonstrate that there is no special or ideal technique and no really clear theories for teaching L2 writing. It is mostly determined by the situation and circumstances, the learners, the curriculum, the teacher, and the classroom objectives. That may be summarized in three key points. First, understanding writing as a complex cognitive task will help in dealing with it as a whole mechanism and not a series of disconnected stages. Second, because language is linked to society, dissociating the relationship between writing and the social environment would undoubtedly cause problems. Third, knowing the history of L2 writers may help in selecting effective and appropriate teaching practices and discarding those that are no longer relevant (Matsuda, 2003). From the same perspective, Kumaravadivelu (2003) states that selecting the optimum teaching approach(es) depends on existing L2 writing knowledge that meets students' demands. Therefore, Horowitz (1986) advises teachers to be cautious when adopting a particular strategy, depending on both academic and practical understanding (Hyland, 2004). We are primarily fostering eclecticism while teaching writing in order to provide students with various opportunities to improve; yet, genre and process focused methods were mostly predominant in the writing syllabus of the current research study. Eclecticism can occur in the current study through selecting activities, techniques, and strategies to offer an opportunity to AfL to be there in the classroom.

1.9.6 Writing Difficulties and Challenges Faced by Algerian EFL Students

Writing is a difficult endeavor for both native and non-native learners since one must be mindful of a variety of criteria ranging from spelling to text arrangement (Rass, 2015). It has been found that Arab EFL learners, in particular, struggle with speaking and writing; specifically, how to use appropriate grammatical rules such as tenses, articles, quantifiers, gerunds, relative clauses, and others, as well as word structure, spelling, and vocabulary (e.g. Diab, 1997; Abi Samra, 2003; Bahloul, 2007). This is due to one major factor: the apparent difference between the two languages, Arabic and English.

Similarly to other students in the Arab world, EFL students in Algeria face significant writing difficulties. According to Ourghi (2002), most first-year university students (80% are low-intermediate and 20% are high-intermediate) are struggling to master basic syntactic structures, writing mechanics, vocabulary, and useful composing strategies. In addition to that, students are suffering from inability to write error-free sentences, grammar problems, language transfer while composing; such as employing some terms from French language (Hamzaoui-Elachachi, 2006), and a lack of cohesiveness and coherence to generate a list of ideas (Hamzaoui-Elachachi, 2010). All of these problems may have an impact on the quality of their work, and thus may impact their achievement in exams. Beside writing-related problems, other psychological issues have been identified, such as students' negative attitudes about writing, low writing self-efficacy, and lack of motivation (Moussaoui, 2012). This is likely to affect their learning self-confidence and self-esteem. Among the solutions proposed to consider for the sake of improving the situation are regular assessment (Moussaoui, 2012) and suitable responsive educational and pedagogical programs (Bouhadiba, 2000).

Our attempt to examine the research field of EFL writing in Algerian context concluded that studies addressing this issue have identified crucial and realistic troubles in student writing. This has significantly alerted researchers to invest in this field. Despite the fact that studies of both exploratory and empirical nature have been undertaken in this area, the focus is mostly on instructional pedagogies as well as learning styles, preferences, and techniques. We could not locate any research that explicitly tackled AfL to assist EFL students improve their writing. This was one of the major reasons to embark on researching this topic.

1.10 Paragraph Writing

We focused on paragraph writing as the core concern because the writing syllabus of the intervention was purely devoted to this type of texts, a paragraph, not essays or compositions. While examining the literature on paragraph writing, we found no research investigating paragraph in the EFL setting. The majority of the discussion regarding paragraph writing occurs in writing textbooks (e.g. Zemach & Islam, 2005a; Oshima & Hogue, 2009). Duncan (2007) explored this gap in the literature to address this issue, revealing that debates on paragraph theory in English composition were common from the mid 1960s to the early 1980s, but that disappeared in the previous two decades. To emphasize its significance, Parks, Levemier, and Hollowell (1986) argue that composing paragraphs is a useful approach to learn basic writing abilities. This explains why mastering paragraph writing is the first step toward learners' progress in writing advanced and complex pieces. In this section, two major issues are discussed: First, what a paragraph is stressing what criteria to have to be judged as such. Second, what a good paragraph is emphasizing what criteria to include to be assessed as such.

1.10.1 What a Paragraph Is

Paragraph is defined in different ways. For Grabe and Kaplan (1996),

> paragraph, as currently used in informational writing, is intended to signal a coherent set of ideas, typically with a main theme and supporting information. Whether or not an explicit topic sentence is provided, readers expect the paragraph, as a unit, to convey a sense of information which can be summarized in a single sentence, and which contributes to the organization of the total text. (p. 353).

This definition focuses on the role of a paragraph and some key features that should be present, such as coherence and topic relevance. Zemach and Islam (2005a), on the other hand, see that

> A paragraph is a group of about 6-12 sentences about one topic. Every sentence in a strong paragraph is about the same topic. All of the sentences explain the writer's main idea (most important idea) about that topic. When the writer wants to write about a new main idea, he/she begins a new paragraph.
>
> A paragraph can give information, tell an opinion, explain something, or even tell a short story. The sentences are arranged logically, so the reader can easily understand what the writer wants to say.
>
> In academic writing, a paragraph has a topic sentence that directly tells the reader the main idea. The other sentences in the paragraph, called supporting sentences, give more information about the topic. They add specific details and explanations. In academic English, the topic sentence is usually (but not always) first or last. (p. 9)

The writers attempted to address the meaning of a paragraph in these definitions, concentrating on its content, structure, form, and genre. The distinction may be due to several factors, including form, content, structure and organization, and purpose. For example, a paragraph can be indented or not, entitled or not, have a topic sentence at the beginning or not, have a conclusion sentence or not, and so on. The reality that is sustained is that the paragraph remains the building block of writings such as essays and compositions; letters, reports, and college topics (Hart & Reinking, 1990), or the essential unit of any piece of written discourse (Duncan, 2007; Grabe & Kaplan, 1996).

Before examining the key and major parts of a paragraph, it is vital to clarify the term 'text'. This term should be defined since it obscures the meaning of the paragraph. In fact, it has several meanings. James (2013) believes that the term text can refer to a unit of written language larger than a sentence, which might be a paragraph, but Connor (1996) believes that cohesion is a quality of the text but coherence is related to discourse. In our case, a text is defined as a written content that is cohesive and coherent; therefore, it can be a paragraph. According to O'Donnell and Paiva (1993) and Zemach and Islam (2005a), the basic components of a paragraph are the topic sentence, supporting details, logical order, connectors, ending sentence, unity, and coherence. On the other hand, Grabe and Kaplan (1996) believe that the topic sentence is not necessary for the paragraph's structure, although O'Donnell and Paiva (1993) believe it is. The concluding sentence, however, is defined as a sentence that summarizes (Reid, 1994; Kemper, Meyer, Van Rys, & Sebranek, 2018), restates (Kemper et al., 2018), gives a solution to a problem, forecasts a circumstance, or makes a suggestion (Reid, 1994).

In our study, we are concerned with a paragraph that is considered as an indented and entitled written block, with a topic sentence that encapsulates the main idea and a series of supporting details developed relevantly to the topic and related appropriately. The whole text is cohesive and relevant to the subject at hand. The paragraph ends with a concluding sentence that restates and summarizes, for the descriptive genre, and gives advice or suggestions, and lessons for the narrative genre.

1.10.2 What a Good Paragraph Is

It is difficult to assess the quality of a given paragraph, but a set of criteria must be met in order to write an effective piece. As previously stated in Section 1.10.1, the references about paragraph writing are largely found commercial in nature and are mostly having no assessment section mostly and if found is brief and not fully explained. Another point to stress is that if any sources are found explaining the elements of successful writing, they are mostly focusing on essay writing. In fact, the same rules can be applied to paragraph writing. Basically, there are three types of qualities: grammatical accuracy, grammatical complexity, and cohesion and coherence. In what follows, we briefly discuss these components.

- **Grammatical Accuracy.** Despite being elusive to have a good written piece, this latter must comply with the standards of English syntax and use, commonly referred to as grammar (Frodesen & Holten, 2003). Following Hinkel (2002), grammatical precision is critical in assessing “normative speaker perceptions of the quality of L2 writing” (p. 182). As a result, improper language use may transmit the content incorrectly, leading to misunderstanding or communication breakdowns. In line with this, Frodesen and Holten (2003) assert that

  > if we view grammar as an essential component of all communication - as a set of linguistic resources from which native and nonnative speakers alike select forms based on appropriateness for meaning, for audience, and for contextual demands - then grammar and writing are inseparable.(p. 57)

Maintaining grammatical precision when writing is undeniable and undebatable and all writing assessment strategies and tools do not neglect this component. This is present in international assessment instruments such as the International English Language Testing System (IELTS) Task 1 writing band descriptors (public version) (British Council, IELTS Australia, and University of Cambridge, 2018), as well as the TOEFL writing score guide (ETS, 2000). As it is important, grammatical accuracy has been extensively explored in several publications, including works by Ferris and Roberts (2001) and Truscott and Hsu (2008).

- **Grammatical Complexity**. It is another feature of a well-written paragraph. Because a paragraph is made up of a series of sentences connected to a single topic (Zemach & Islam, 2005a), a sentence plays an important part in the construction of a good paragraph. If a paragraph exclusively contains simple sentences, it does not sound effective since simple sentences can be joined to make compound, complex, or compound-complex sentences. The more complicated the sentences, the more engaging the paragraph sounds and meets the grammatical complexity. Grammatical complexity is also highlighted in L2 writing since it is present in assessment tools such as the rubric by Jacobs et al. (1981), and it has been extensively explored by numerous writers (e.g. Storch, 2009).
- **Cohesion and Coherence**. They are among the most important characteristics that distinguish a written piece from another. Cohesion is based on grammar or meaning, but coherence refers to the reasonable relationship between the ideas in a written piece (Halliday & Hasan, 1976). Coherence is also seen primarily related to content (James, 2013) and can be achieved through using connectors, linking one sentence to another (Wyrick, 1999) to generate meaning for reader (Lee, 2002). An important research was undertaken by Hamp-Lyons and Reed (1990) studying the creation of multiple-trait writing assessment for the University of Michigan’s undergraduate admission assessment to highlight the relevance of cohesion and coherence. According to the data gathered from instructors responding to the definition of excellent paragraph based on how they respond to students’ written text, the quality of the paragraph should entail strong

presentation of "content, argument, text structure features (cohesion), evidence of planning (coherence), and language" (p. 259). In fact, cohesion and coherence are incorporated in all assessment methods as in the IELTS and TOEFL, and they have been widely examined in L2 contexts (e.g. Simpson, 2000; Liu & Braine, 2005).

To summarize, assessing a given paragraph, conforming to the specified criteria and standards, requires being articulated in content and interrelated in ideas, developed relevantly to the topic, devoid of grammar errors, and meaning contextualized.

Part IV. Empirical Studies

The major aim headed within this section is reviewing the most relevant studies related to the current research. Before starting reviewing them, we see that it is useful to pinpoint that works that accurately represent the study under investigation seem inexistent and the studies that are tightly relate to AfL in EFL writing context are scarce. Therefore the studies to be reviewed here might interplay to shape the dimensions of the current study. In other words, this research study is not a replication of a specific, previous study. Regarding the lack of studies, Lee (2017) and Lam (2018) claim that there are relatively few studies in the field of examining EFL writing under AfL orientations and implementations. This scarcity is explained by the recent movement to adopting AfL, particularly in EFL writing setting, which takes time since changes begin with theoretical attempts and considerations before moving to practical concerns. For further clarification, Hamilton and Pinnegar (1998) explain that shifting behaviors arise from a higher degree of appreciation for the changes after being aware of and comprehending the theoretical ones. Regardless of accepting and understanding the change to proceeding for change on terrain, in this field we can add the difficulty and complexity of the skills themselves, writing and assessment, which might be seen as another hindrance for advancing research and producing abundant research papers. To make it clear, one can find many studies that treat concepts such as self-, PA and rubrics, but the scarcity we mean here is that those studies consider the aforementioned concepts as activities and tasks in classroom without taking assessment as an integral part of classrooms.

Before beginning the actual reviewing, some criteria were pre-set in order to provide an accurate examination of the research studies chosen. To begin, the study should be in the framework of EFL. Second, the intended level is university. Third, the target skill is EFL writing. Fourth, the search date was different, including the current year, up to July 2022, to keep the review up to date. Fifth, Google Scholar was utilized for this purpose since it has a large number of references and has access to many other databases such as ERIC, Web of Science, and PsycNet. The result of this search, following the aforementioned criteria, ended up with selecting six studies as the most relevant and appropriate.

To best present the review, we see that it is practical to use a table in which we display the research problems formulated as research question(s) and/or hypothesis (es), and directly below it we provide and briefly describe the primary findings of each study.

Table 1.7 *Empirical Studies Summary*

Author(s)	**Research problem**	**Research questions and/or hypothesis**	**Research design and methods**
Wang (2016)	The primary research goal of this study was to examine students' perceptions of their usage of instructional rubrics, i.e. for formative purposes, in their EFL writing and to identify the characteristics that catalyze its efficacy.	Two research questions were developed for this research study: 1. How did students perceive the rubric's role in self-assessment, especially in relation to their self-regulated learning of writing? 2. What factors, if any, were perceived by the students as affecting the rubric's effectiveness in self-assessment in the writing class?	This research included 80 students (24 male and 56 female) chosen from one of China's universities. The data was gathered via reflective diaries and six case studies (2 male and 4 female) who were interviewed retrospectively.

Belachew, Getinet and Gashaye (2014)	The purpose of this study was to investigate EFL teachers' perceptions and students 'practices of self-assessment in the writing classroom at one of Ethiopia's universities.	The following research questions were addressed in this study: 1. How do EFL students perceive self-assessment? 2. What kinds of perceptions do EFL writing teachers have toward self-assessment? 3. Do students assess themselves genuinely and rate their level in writing effectively? 4. What does the experience of EFL writing teachers toward self-assessment look like?	The participants selected for this study were 50 University students and 10 teachers. The authors used mixed method, combining statistical outcomes and qualitative results to respond to the research question formulated. Semi-structured interviews, document analysis, and focus-group discussions were used to collect data.
Wang (2013)	The study's goals were to look into how students' perceptions of their peers' feedback on their EFL writing changed over time, to look into the factors influencing their perceived usefulness of peer feedback for draft revision, and to look into their attitudes toward the use of a rubric in peer feedback practice.	This study aims at answering the following research questions: 1. Do Chinese students' perceptions of peer feedback on their EFL writing change over time? 2. What factors may affect students' perception of using peer feedback for draft revision? 3. How do students perceive the rubric's role in their peer feedback practice?	The study included 53 Chinese EFL students, as well as six case study informants. Questionnaires, interviews, and students' reflective writings were used to collect data.
Birjandi and	The study was conducted in an attempt to compare the effectiveness of various assessment techniques,	The following research questions were addressed in this study: 1. Does journal writing as a self-assessment practice improve participants' writing performance, as	The study was designed as a quasi-experimental, pre-test-post-test design. The participants included in this study were 157 intermediate

	namely self-assessment, PA and TA, in promoting Iranian EFL learners' writing performance.	compared with TA? 2. Does self-assessment improve participants' writing performance, as compared with TA? 3. Does PA improve participants' writing performance, as compared with TA? 4. Does self-assessment accompanied by PA improve participants' writing performance, as compared with TA?	EFL students, mostly were girls. They were divided into five groups, four experimental groups and one control group, to undergo each a treatment. The first experimental group used journal writing as a self-assessment approach, the second used self-assessment, the third used PA, and the fourth used both self- and PA. TA was found in all experimental groups except the fourth, which performed both self- and PA. Only TA was used in the control group. Several tests were employed(pre-test and post-test and the Jacobs et al. (1981) rating scale to assess students' works.
Diab and Balaa (2011)	The purpose of this study was to create rubrics to assist students improve their writing. Following that, they were instructed to write criticisms based on those rubrics, and they revised what they had written in response to their teacher's grades and remarks. After that,	The authors of this study hypothesized that using rubrics helps students improve their writing drafts effectively.	The participants were 78 undergraduate students enrolled in three sections at the American University in Lebanon. The data gathering tools used were an online survey to yield their opinions concerning rubrics' use and a T test was used to explore the effectiveness of rubrics on students' writing performance.

	students were questioned on the usage of writing rubrics		
Lam and Lee (2009)	The study was carried out in order to explore how portfolio assessment can be exploited formatively in EFL writing classroom.	Three research questions were developed for this research study: 1. What are students' and teachers' perceptions of the impact of the portfolio process on student writing? 2. What are students' and teachers' views of the summative and formative functions of portfolio assessment? 3. How can portfolio assessment be utilized to realize the formative functions of writing assessment?	The present study's participants were enrolled in a sub-degree program at one of Hong Kong's universities. They had different English levels, ranging from low intermediate to intermediate, and the majority of them had a low level. Student questionnaires and student and teacher interviews were employed to collect data.

Study by Wang (2016). The findings of this study revealed that rubrics may be used effectively for formative purposes in fostering self-regulated learning through four stages: goal-setting, planning, self-monitoring, and self-reflection. Concerning the factors mediating the efficiency of rubrics, the author discussed the results in light of the theoretical model proposed by Panadero and Jonsson (2013). The authors examined how to improve students' self-efficacy, to help the feedback process as well as the factors such as intervention time, educational level and gender, and to moderate rubrics' learning impacts. Following that model, Wang (2016) classified factors in student self-assessment performance as those related to rubrics and rubric users. They were category and structure coverage, quality level descriptors and score range, rubric users' knowledge domain, and intervention length.

For rubric coverage and structure, Wang believes that in order to ensure effective self-assessment, students must be effectively involved and given enough space to decide on rubric content and structure. Another significant aspect he brought to light was students' perceptions of analytic rubrics. He observed that not all students welcomed the analytic structure of the rubric to self-assess their writing, and he suggested being flexible in allowing students to choose the appropriate rubrics structure to use. For quality descriptors and score range, the author suggests selecting acceptable wording in rubrics using plain and simple language, and saw that complementing rubrics by scoring range would add greater clarity to the description levels. Concerning rubric users' domain knowledge, the author above proposes developing explicit underlying goals for the rubrics produced and taking into account students' prior experience, as he observed with his participants. Finally, the author took past research into consideration, such as that done by Andrade, Du, and Mycek (2010), and discovered that students required to be familiarized with the criteria of rubrics two to three times. Furthermore, he warned against using solely general-task rubrics for an extended length of time and proposed augmenting them with a specific-task kind.

Study by Belachew, Getinet and Gashaye (2014). One of the major findings reported in this study is that both participants and teachers were enthusiastic about self-assessment. Over the first four sessions in which students were required to self-rate their works, the researchers reported a decrease in overrating. The participants were divided into three groups by the researchers: overrated, underrated, and really rated. Overrated percentages were 56%, 50%, 46%, 44% while the genuinely rated percentages were respectively 30%, 32%, 38%, 40%. For each session, the under-rated represent the outcomes of 100%- (Overrated + Genuinely rated). The researchers discovered that teachers had basically no experience with self-assessment and believed that students could not grade their papers. However, they noticed that self-assessment was intriguing since it helped students be independent and acquire cognitive reflective habits in order to accomplish good work.

Study by Wang (2013). The findings yielded show that students' perception of the usefulness of writing peer feedback decreased over time. The author's study contradicts Mangelsdorf's (1992) results, which suggest that students' belief in the efficacy of peer feedback practice would increase with time. The five factors affecting this decrease were: (a) students' knowledge of the topic assigned to essay writing, (b) students' limited English proficiency, (c) students' attitudes toward peer feedback practice, (d) time constraints of the in-class peer feedback session, and (e) students' concerns with interpersonal relationship. Furthermore, the students perceived the rubric as an explicit assessment guide to evaluating their peers' EFL writing but also perceived negatively. This latter point was explained by rubric's overuse, which resulted in students doubts of its applicability.

Study by Birjandi and Tamjid (2012). The findings of this study are diverse and all statistical in nature. There was no significant difference in writing scores between the groups that used journal writing with TA and those that solely used TA for the first study question. For the second research question, there was a statistically significant difference between the groups that used TA and self-assessment and those that solely used TA. Similarly, a statistical significance was marked between the groups performing PA and TA and the group using simply TA for the third research question. However, no statistical significance was established between the self-and PA groups and the TA groups for the fourth study question**.**

Study by Diab and Balaa (2011). The findings of this study were organized into two types: (a) student performance on the first and second versions of the writing critiques, and (b) student opinions about the usage of the rubrics. The T test revealed a statistical significance between students' grades on the first and second drafts, and the authors claimed that the use of rubrics was not the cause of this significance in grades; nevertheless, this result suggested that it might help students revise their drafts more effectively.

To back this argument, the authors used survey results that demonstrated good attitudes and reactions to the usage of rubrics, with 95% agreeing that the rubrics helped them improve their initial draft. Furthermore, some statistical results concerning rubrics have been reported, such as 80% of students saying that they used the rubric for the first time, 97% confessing that the rubric helped them understand the assignment and how the teacher would grade them, 82% saying that the rubric helped them understand their weaknesses and strengths in writing, 91% agreeing that grading using the rubric was fair and rubrics were concise and clear, and 100% feeling satisfied with the time spent introducing the critique rubrics.

Study by Lam and Lee (2009). The findings of this study highlighted three main benefits of using portfolios: 1) influencing student motivation for writing using portfolio assessment: (a) autonomy in selecting their best work, (b) improving their writing through conferences (such as providing verbal feedback, and (c) providing a suitable learning environment to enhance their writing. When asked if the portfolio approach had the capacity to improve their writing performance, the majority of students (83.3%) replied positively, with two significant benefits reported: improvement in accuracy and idea generation.

Through the portfolio process, students become more aware of the distinctions between the typical product-oriented and portfolio-based classroom. To offset the assessment portfolio's summative emphasis, the researchers opted to emphasize teacher scaffolding step for (a) continuing teacher feedback, (b) conferencing, and (c) peer reviewing. Students understood that the assessment atmosphere in the portfolio program was less threatening and more supportive than time-constrained and unscheduled essay tests after receiving formative input from tutors and peers. Despite their positive responses to the formative parts of portfolio assessment, students preferred summative grading because they feel that grades better inform their writing standards.

1.11 Chapter Summary and Conclusion

AfL led fundamental change in language teaching and learning because it represents a shift from focusing solely on teaching and learning materials and methodologies to incorporating assessment into the classroom and, more significantly, placing the student at the center of attention. This is a growing focus on the terrain, particularly in the context of EFL writing, gaining increasing interest despite the intricacy of both aspects; assessment and writing. Worth bearing in mind is that AfL and FA are closely linked, with some debate over whether they should be used interchangeably. Furthermore, FA is included in the discussion, whether or not it is associated with SA.

In fact, AfL is a concept that may be operationalized in different ways depending on the aims and objectives for which it is employed. It effectively incorporates feedback, student assessment (self- and PA), and teacher scaffolding and guidance with careful execution. These areas are all hotly debatable. To perform assessment within AfL, some tools, namely rubrics are suggested for implementation to be used to help students grasp the assessment criteria and procedures so that they may participate effectively in assessment performance. AfL may be motivating and supporting self-regulation learning, autonomy, and lifelong learning because it primarily aims at further improvement. For these purposes, AfL is not fully embedded within a specific theory and, in fact, lacks its own theory. Finally worth mention is that scholars are still working on developing an adequate theoretical niche for this concept.

2

CHAPTER TWO: RESEARCH DESIGN AND METHODS

After reviewing the major concepts, discussing the theoretical background, and identifying the gap in the literature toward the need of investigating AfL to help student improve their EFL writing, in this chapter the blueprint of the study conducted is presented accordingly. This chapter gives an account of the objectives highlighted and presents the research questions formulated in addition to the context of the study, including population, sampling and sample, and setting. A section about the study nature, research design and methods, triangulation, and validity and reliability of the study is fully documented. Following that, the data gathering tools employed are thoroughly described in terms of content, and piloting and administering phases. The last section of this chapter is devoted to data gathering procedures and analysis.

2.1 Research Questions and Objectives

Taking EFL participants' needs and requirements in writing into consideration and regarding the various relevant aspects discussed in the literature review chapter, the current study attempts to achieve the following major aims:

- explore first year EFL students' difficulties in writing;
- test AfL implementation procedures' effectiveness; including writing rubric use, student assessment, teacher assistance, and feedback;
- engage students in an interactive classroom with their teacher and peers based on constructive feedback; and
- engage students in self-regulation learning for further motivation, self-esteem, self-confidence, autonomy, and lifelong learning as long term goals.

In view of the objectives set for this study, the research questions formulated are put forward:

1. What writing difficulties might be highlighted in first year EFL degree students?

2. Does the use of writing rubrics help first year EFL degree students overcome difficulties in writing?
3. Does self-assessment help first year EFL degree students overcome difficulties in writing?
4. Does anonymous peer assessment help first year EFL degree students overcome difficulties in writing?
5. Does assessment for learning help first year EFL degree students improve their writing?

The research questions formulated are hybrid, or mixed and integrated, as they convey the content of the study, which is based on mixed methods (Tashakkori & Creswell, 2007; Creswell & Plano Clark, 2011; Creswell, 2009/2014). The first research question examines the existence of problems in EFL writings and the subsequent questions all flow into examining whether AfL procedures, as operationalized in this study, are effective in helping students progress in writing and whether they are going not to prove or confirm a given theory but to probe it (Creswell, 2014).

2.2 Context of the Study

It is commonly known that a research study has no influence unless it is immersed in a context; thus, a researcher must explicitly present the design and conception of the study (Gersten, Fuchs, Compton, Coyne, Greenwood, & Innocenti, 2005). Via aligning the reasons for adopting a quasi-experimental design as explained by Seliger and Shohamy (1989) to the study at hand and discarding any possibility of conducting true experiments in education (Griffee, 2018), the in-practice reasons are threefold: (1) The current research was conducted under natural conditions in an educational setting at university; therefore, it would meet the external validity criterion. (2) The second reason is getting an easy access to the target population and avoiding any reluctance from the participants' part to participate in the study, as it might happen in the true experimental research. (3) It is less "disruptive and intrusive" (p. 149) to conduct quasi-experiment i.e. any limitations from the administration part were avoided such as the need to advance in the program because the study was in accordance with the main objectives of the writing syllabus. To contextualize this study, three major components are included: the population, sampling and sample, and the setting.

2.2.1 Population

Population is defined as the entire realm of people (or items) that can be measured or counted (Knapp, 2017). The population of this study consisted of all first year EFL students enrolling at the department of English language at the University of Algiers 2 during the academic year 2018-2019. The total number of students in the population was estimated to be 943 students following the statistics provided by the administration.

The researcher worked with first year students other than the other levels for various reasons. First, providing feedback on students' academic performance matters particularly for first year students to build in students a solid platform (McInness & James, 1995). This is one of AfL's primary roles and purposes. Second, the earlier students' experience, the more values, attitudes, and methods to learning they develop, which they will face throughout their tertiary experience (McInnis & James, 1995). In line with this, as writing is a challenging and crucial skill for students, teaching them the fundamentals can help them go to more advanced levels with less difficulty. Third, engaging students effectively with learning stems from the quality and the nature of their first year university experience (Krause et al., 2005; Kift & Nelson, 2005). Students may understand how they will handle the coming years and how significant the challenges will be if they are trained to identify writing difficulties early, in first year, and engaged in working on their problems. Fourth, if students understand what is expected in the assessment activities and resources through constructive guidance and feedback, they are likely both to enhance their learning and establish a solid foundation to understand how to approach the writing skill in the subsequent years (Hendry, Armstrong, & Bromberger, 2011). As a result, all that may help students acquire skills such as evaluation, analysis and synthesis, understand the importance of feedback, and identify and apply the assessment criteria.

2.2.2 Sampling and Sample

Including the whole population in research happens only in census, but in other circumstances this is unnecessary and would be a waste of resources (Dörnyei &Taguchi, 2010). Nunan (1992) states that getting data from the complete population is impossible, and selecting a representative sample is essential. Using a sample rather than the entire population is faster, simpler, and less expensive (Dörnyei & Taguchi, 2010; Knapp, 2017), and if done correctly, the results can give quality information on the entire population (Knapp, 2017).

The sample is defined as a "subset of individuals from a given population" (Knapp, 2017, p.27). The sample of this study consists of two groups attributed randomly by the administration. Table 2.1 displays the number of participants in each group.

Table 2.1 *Participants' Number in Experiment and Control Group*

Participants	Experiment group (or Intervention group) (Student number)	Control group (Student number)
The number listed by the administration	63	62
Spending just one semester	3	/
Not regular for various reasons: working, personal problems, and others	3	4
Total absence for reasons known or unknown to the administration	17	17
Left after the 1st semester's results	2	2
Participants	38	39

Following Dörnyei and Taguchi (2010), good sample hires similar characteristics of the population; the often quoted "the larger, the better" principle, alone, is unhelpful for researchers (p. 62). It seems that there are no pre-set and fixed rules to follow to set an optimal sample because every research study is unique (Nunan, 1992; Dörnyei & Taguchi, 2010; Knapp, 2017). The study sample was chosen in accordance with the major objectives of this research, the highlighted aims, and the type of the constructs to be dealt with during the intervention. In terms of overall population characteristics, the study sample had comparable features in terms of age range, linguistic levels (mixed skills; overall low pre-intermediate), and gender number predominance; females outnumber males in all groups.

In fact, sampling bias can occur unintentionally and that may threaten the external validity of the findings. A very small sample causes instability of the statistical findings wherein the large ones are difficult to handle and process because they are time, energy and, may be, money-consuming as well as the possibility of placing additional subjects at potential risk (Knapp, 2017). The sample selected for this study was cared of to be handled by the researcher, responding to the requirement of the research planned. This was in agreement with the objectives of the study, the nature of the dependent and independent variable(s), and the practicality of the intervention.

Having a control group is a crucial issue to shed light on. This provides a more realistic evaluation by comparing the outcomes of the intervention group to the control one (Knapp, 2017). Despite having no possibility to randomly assign the participants into groups in this study, one intact group has been selected as the intervention group and the second one as control group based on the outcomes of the pre-test. Both groups were assigned by the administration; this means that the participants' number, gender, and students' levels in writing was not known beforehand.

2.2.3 Setting

The intervention took place at the department of English in the University of Algiers 2, Abou El-Kacem Saadallah. There were two key reasons for selecting this department. The first was that the researcher completed her doctoral research and was permanently supervised at this university. The second reason was that the research objectives went in agreement with the content of the department's writing syllabus.

2.3 Research Methodology

The present section includes an account of the research methodology that builds the practical occurrence of this study. It includes six sub-sections: the study nature, characterizing the variables of the study, research design, research methods, triangulation, and validity and reliability of the study.

2.3.1 Study Nature

Posivistic approach and pure quantitative method are largely rejected in researching writing, even though in some contexts they would be appropriate, because in this field “qualitative, natural, and ‘thicker’ data collection techniques” are believed more appropriate (Hyland, 2009, p.151). The current study is quasi-experimental in nature and design, called exclusively experimental as well because it took place in an educational setting and includes three basic components: population, the treatment, and the measurement of the treatment (Seliger & Shohamy, 1989). It is deductive oriented research; theory happens before collecting data (Griffee, 2018), or theory guides research (Bryman, 2012) and the researcher conducting this type of research believes the existence of legitimate objectives to accomplish at the end of the research process, relying on the power of experimentation to be the sole tangible reason to verify or disprove the effectiveness of a specific plan (Griffee, 2018).

A quasi-experimental design is characterized by the implementation of an experimental procedure without having full control over all extraneous variables and no random assignment of participants to groups as in true experimental designs, yet it is nevertheless “better at controlling extraneous variables than the weak experimental designs” (Christensen, Johnson, &Turner, 2014, p. 270). Following the dimensions and major categories of social sciences proposed by Neuman (2014), this study is an applied research that comes under the type of evaluation studies. This category comprises studies that address the following question: “Is a program, a new method of doing things, a marketing effort, a policy, and so on effective?” (p. 28). In agreement, the following question may be formulated to describe this study: Does AfL assist EFL students enhance their writing performance?

Second, beyond the ordinary dichotomy basic vs. applied, this study provides both instrumental, “task-oriented actions and principled, value-based, engaged behaviour”, and reflexive knowledge, “self-aware, value-oriented knowledge” (Neuman, 2014, p.

34). Providing instrumental knowledge in this study can be highlighted in student assessment performances, self- and APA, and the use of rubrics and reflexive knowledge can be deduced from the meaning of AfL for purpose of motivating and promoting, and the importance of the constructive and elaborated feedback. The primary goal of conducting this study is to raise participants' awareness about the need of taking care of their own writing progress from their early first year of university. Furthermore, it attempts to self-regulate students' skills in order to equip them with strategies to face their deficiencies and capitalize on their strengths in order to be autonomous and lifelong learners.

Third, following Neuman (2014), this study is explanatory based on descriptive and exploratory research. It is explanatory because it tests hypotheses generated from theory. Based on exploratory and descriptive grounds, the rationale for conducting this research was fully explained in the general introduction and literature review chapter. The next stage is known as the confirmatory stage; it is where an intervention is carried out. The participants acted in classroom assessment to reflect on the assessment findings using rubrics as assessment tools and practice student assessment, self- and APA, under teacher scaffolding instruction to provide effective and detailed feedback in order to advance in their writing.

2.3.2 Characterizing the Dependent and Independent Variables

It is critical to identify and treat variables in experimental design because this may not be the case in other designs (Griffee, 2018). The researcher tried to characterize the variables using various researchers' perspectives in order to conceptualize the variables and to construct the study at hand.

First, for Collins, Joseph, and Bielaczyc (2004) three types of dependent variables are important to assess: (a) climate variables such as engagement, cooperation, risk taking, student control; (b) learning variables, such as content knowledge, skills, dispositions, metacognitive strategies, learning strategies; and (c) systemic variables, such as sustainability, spread, scalability, ease of adoption, and costs (p. 36). In this study, the climate variable is the students' control of their accuracy and fluency in writing at individual level and the degree of responsiveness when performing self- and APA. The learning variable is the students' progress in writing, and the systemic ones are self-progress, motivation, self-esteem, and autonomy

Second, for Bordens and Abbott (2018) dependent variables can be behavioral measures, physiological measures, self-report measures, and implicit measures. Excluding physiological measures from this study, the behavioral measures include measurements like "number of errors made, number of responses" (p. 160). Self-report refers to the ability of participants to report on their own behavior and can be "prospective (speculate on future behavior) or retrospective (report on past behavior)" (p. 160). For self-report in this study, self- and APA were accompanied with specific-genre writing rubrics with follow-up sections in which the participants can report on both their writing progress and their use of rubrics. Implicit measures assess "unconscious reactions to stimuli and are used to tap into attitudes that individuals may not admit to overtly" (p. 160). For implicit measures, the researcher tried to explore what students had acquired at the end of the intervention in relation to the overall improvement, such as self-confidence and autonomy in writing.

Last but not least, following Knapp (2017), in order to select, document, and run the results of statistical tests, it is essential to know about the type of variables involved in data set or research design. The participants' quantitative measures are continuous ratio variables, whereas the qualitative assignments are categorical ordinal. In this study, the written pieces are rated as categorical ordinal, such as poor, average, good, or excellent (These labels are used because they are widely known), and the scores gathered are represented by a numerical interval, with some constraints that are explicitly stated in Section 2.4.3.4- The scores might range from 0 to 20. Combining interval and ordinal scales in measurements provides less information if compared to interval and ratio scales; thus, using interval and ratio scales should be used whenever possible (Bordens & Abott, 2018).

2.3.3 Research Design

The importance of research design in any research is that it maps the occurrence of a particular study to fulfill certain objectives. It is seen as a road map for how to thoroughly undertake a certain study in order to best accomplish specified objectives (Tashakkori & Teddlie, 2003). Creswell, Clark, Gutmann, and Hanson (2003) define design as a procedure for gathering, interpreting, and reporting research. In other words, research design, according to Griffee (2018), is an operational model or blueprint for a research effort that caters for "internal reasoning (causality) and external reasoning (generalizability)" (p. 22). In fact, each study design is determined by its particular implementation which varies widely depending on the participants' requirements, interests, skills, interpretations, interactions, and aims (Collins et al., 2004).

The significance of research design is that it shapes the frameworks for data collection and analysis, reflecting a set of decision-making procedures for prioritizing a variety of factors in research, including:

1. expressing causal connections between variables,
2. generalizing to larger groups of individuals than those actually forming part of the investigation,
3. understanding behaviour and the meaning of that behaviour in its specific social context, and
4. having a temporal (that is, over time) appreciation of social phenomena and their interconnections (Bryman, 2012, p. 46)

It is worth emphasizing that quasi-experimental research occurs between pre- and true experiments, and the major feature is that there is no complete control over variables since the groups are not randomized, i.e. they are pre-assigned (Seliger & Shohamy,1989; Nunan, 1992; Griffee, 2018; Salkind, 2018). According to Bordens and Abott (2018), pure quasi-experimental design is similar to true experimental design with the exception of using a quasi-independent variable instead of the true independent one. For recall, the researcher of this study had previously worked with pre-assigned groups by the administration (as explained in Section 2.2.2 in this chapter). The most desirable aspects of any good research design are the random selection of participants and the use of a control group. The former is the preferable design for obtaining an accurate and impartial evaluation of the effect of an intervention (Shadish, Cook, & Campbell, 2002; Bordens & Abotts, 2018), while the latter serves as a basis for appraising the overall treatment. Because true experiments are thought to be hard to conduct in educational contexts, different experimental designs should be considered. This study employed a non-equivalent comparison group design to examine the effectiveness of AfL implementation and whether it may improve EFL students' writing.

Participants were divided into two groups. To avoid pitfalls that may result from participant selection or non-selection, introducing a control group is essential to avoid any problems with internal validity (Bordens & Abbott, 2018), or /and this may also correct the internal validity, if affected, which would appear due to the quasi-nature of the experiment (Nunan, 1992).

2.3.4 Research Methods

Many methodologies, strategies, processes, and techniques may be used to approach researching writing, "including quantitative as well as qualitative methods and approaches" (Hyland, 2009, p. 141). The mixed method is the most recent and widely used term to describe a mixture of quantitative and qualitative methodologies (Creswell, 2014). It entails "combining both qualitative and quantitative data gathering and analysis methodologies in a single study" (Creswell, 1999, p. 165). This is extremely useful, and it is well acknowledged and valued by scholars (Creswell et al., 2003; Creswell, 2009). Overall, the most convincing experimental evidence comes from triangulating research designs and measurements (Gottman, McFall, & Barnett, 1969, p. 299).

The method adopted in this research is the mixed one and the design is based on both theory and practice (Collins et al., 2004). Following Creswell (2014), the quantitative data targets marking whilst the qualitative data describe the processes experienced by the participants in the treatment groups. Furthermore, Spada (1990) sees that the qualitative data are often necessary to accompany the interpretation of the quantitative ones. Another major reason to use a hybrid method is to boost research credibility. According to Feuer, Towne, and Shavelson (2002), if a study conjecture or hypothesis can survive inspection by numerous approaches, its credibility is considerably boosted, and therefore the entire research might be reinforced (Greene & Caracelli, 1997). On the other hand, combining qualitative and quantitative approaches covers the flaws of each side, neutralizes them, or may even cancel out some of the drawbacks (Jick, 1979). Last but not the least, following Needleman and Needleman (1996) qualitative approach in intervention research can complement the quantitative study or stand as separate, but the purpose in both cases is to gain an in-depth understanding of the process under study. That can be in four principal ways:

1. exploring the context setting before the intervention proper
2. checking the reality to correct any oversimplification in the quantitative methodologies
3. helping to simplify and standardize complex social phenomena in aims of quantitative analysis
4. using qualitative data to describe and explain the study conducted

2.3.5 Triangulation

To trace the history of triangulation in the literature, it first appeared in Campbell and Fiske’s (1959) paper where they explored the convergent and divergent validation of measuring tools. It was further expanded upon in Webb, Campbell, Schwartz, and Sechrest’s (1966) explanation of ‘unobtrusive measures’. Later, it was introduced into discussions of qualitative methods by Denzin (1970). Its value has gradually grown among scholars, and it is now important in the discussion of mixed methods and mixed strategy research (see, for example, Erzberger, & Kelle, 2003). Basically, triangulation came to acknowledge that no measure in social science is perfect, and measuring the construct just one time is subject to error and biases (Heath, 2015).

The word triangulation, just as mixed methods, has been widely used (Hammersley, 2007). It was created by quantitative researchers, but it has lately become a technical term used by qualitative researchers and has become a major problem in mixed method research (Hammersley, 2007; Bergman, 2008). The meaning of this term, triangulation, seems clear and universally accepted, yet there are different inquiries about it (Bergman, 2008). It is described as a “magical word” (Tashakkori & Teddlie, 2003, p. 674) and “near-talismanic method approach” to integrate qualitative and quantitative research (Miles & Huberman, 1994, p. 266). On the other hand, it has been claimed to be “over-used to the point where it means nothing” (Tashakkori & Teddlie, 2003, p. 674) and to “have too much meaning,” therefore having “no meaning at all” (Sandelowski & Barroso, 2003, p. 328).

Clearly, this term does not have a single definition (Bergman, 2008), but rather a variety of, at times contradictory, conceptualizations. For Greene et al. (1989), triangulation is a mixture of two or more approaches used independently and concurrently to analyze the same conceptual phenomena. However, Bergman (2008) has tried to explain triangulation in different ways, including validity checking, indefinite triangulation, finding complementary knowledge, and epistemological dialogic or juxtaposition. In terms of validity checking, triangulation is described to be closely tied to the aims highlighted regardless of combining quantitative and qualitative methodologies; nonetheless, the latter is always legitimate and acceptable if the researcher opts for it. Indefinite triangulation, the second interpretation, explains how triangulation is used to gather divergent interpretations at the expense of validity of inferences from data. To hunt for complementary information, the well-known meaning of triangulation, it is ensured via employing diverse sources to collect data. Bergman (2008) explains triangulation as epistemological conversation or juxtaposition following Flick’s (1992/1998) conceptualization, who noted that triangulation not only collects information using multiple means but also shapes the world in distinct ways.

As it has various definitions and sometimes contradictory perceptions and explanations, the researcher of this study sees that triangulation is the use of multiple data sources, theories, or methods (Patton, 1990) to answer the research questions appropriately because that can bring greater plausibility to the interpretation of results (Hyland, 2009). This is supported by two major arguments, both of which are related to mixed methods: (a) mixed methods are dedicated to the design of instruments as well as the transformation and triangulation of data (Borkan, 2004), and (b) mixing methods are used to provide an in-depth description of the participants in research study using the qualitative and quantitative methods, by exploring and investigating various facets. This can also help with generalizability and statistical dependability, if any (Creswell et al., 2003).

2.3.6 Validity and Reliability of the Study

In this section, the researcher discusses the validity and reliability of this study. Starting with validity, the validity of any study is a sine qua non concern for making a particular research ethical and acceptable. In fact, various types of validity are found, and each or a mix of two or more, is defined by the study conducted. According to McCleary, McDowall, and Bartos (2017), the most crucial kind is internal validity since losing control over it can impact and harm representativeness and generalizability, if looked for (Campbell, 1957). This type is more connected to external validity in terms of generalizability, which involves generalizing the results of an experiment to multiple participants, contexts, settings, and factors (Campbell, 1957; McCleary et al., 2017). As a matter of fact, generalizability is not our primary focus in this study because we examined a new issue in the context chosen, Algerian higher education; therefore, achieving the goal of representativeness undoubtedly needs more research. The primary goal of internal validity in this study is to show that the explanations of a particular occurrence, topic, or collection of data provided by a piece of research may really be maintained by the data (Cohen, Manion, & Morrison, 2007). More importantly, as in this study, quasi-experimental designs have a greater degree of internal validity than pre-experimental designs with no control group, and they can have substantial external validity, possibly as high as real experimental designs (Salkind, 2018).

If internal validity is viewed as discrete threats (Campbell & Stanley, 1966; Cook & Campbell, 1979; McCleary et al., 2017), in this study it is basically linked to instrumentation and selection. To maximize validity of instrumentation in addition to designing valid and reliable tools, being also piloted as a crucial step to maximize validity, the researcher opted for the mixed methods as this latter offers particular opportunities for increasing the validity of a study through: (a) triangulation (i.e., seeking convergence and corroboration of findings from different methods that study the same phenomenon); (b) complementarity (i.e., seeking elaboration, illustration, enhancement, and clarification of the results from one method with results from the other method); (c) initiation (i.e., discovering paradoxes and contradictions that lead to a re-framing of the research question/questions); (d) development (i.e., using the results from one method to help inform the other method); and (e) expansion (i.e., seeking to expand the breadth and range of the investigation by using different methods for different inquiry components) (Greene, Caracelli & Graham, 1989). In fact, validity is viewed differently by both quantitative and qualitative traditions, but it is agreed that the link is established via connecting the procedures of collecting data and making inferences from those data (Maxwell & Mittapalli, 2010). On the other hand, achieving validity in selection is viewed as using a design with a control group since this is more powerful than having none (Nunan, 1992; Bordens & Abbott, 2018).

As for reliability, this dimension is inextricably linked to the quantitative heritage. If a study is to be considered reliable, it must demonstrate that if it can be repeated on a similar group in a similar situation, identical results would be obtained (Cohen et al., 2007). However, reliability in qualitative paradigm is often discussed and replaced with terms like dependability (Lincoln & Guba, 1985) or trustworthiness (Shenton, 2004). When it comes to empirical results, reliability becomes a questionable issue (Campbell, 1957). It is pointed out to that non-laboratory findings look untrustworthy, but they may be readily extrapolated to different contexts, whereas laboratory results appear trustworthy, but they are difficult to generalize. As a result, empirical findings might be either trustworthy or generalizable, but not both (McCleary et al., 2017). To ensure the reliability of this study for further replicablility, as this latter is tightly connected to reliability (Bryman, 2012), the researcher of this study maximized that through designing and piloting carefully the research tools in addition to collecting the data needed diligently, and analyzing and interpreting them conveniently. Reliability looked for behind this study is related to AfL procedures' implementation. In other words, the extent to which other researchers may reintroduce the same intervention following the same steps and guidelines.

2.4 Data Gathering Tools

Researching writing tends primarily to prefer data gathered in naturalistic rather than under controlled conditions. It differs from more quantitative research in the way of viewing reality, the relationship established between the researcher and the research subject, generalizability issues, and causality. While data elicited through experiments, questionnaires, and structured interviews can be useful, data gathered via analysis of authentic texts and observations are seen more common (Hyland, 2009, p.144). This section provides an overview of the rationale, design, validity, and reliability concerns, as well as the description and goals of each of the instruments used to elicit data from informants. Pre- and post-tests, pre- and post-study questionnaires, rubrics and follow-up sections, the working portfolio, and the researcher journal report are the instruments in question.

2.4.1 Pre-test and Post-test

Pre-tests and post-tests can be used to evaluate three kinds of learning variables "content, reasoning, and dispositions" (Collins et al., 2004, p. 36). The dependent variable of this study is EFL writing which is a content one. Because the researcher must ensure that the pre-test is appropriate for the participants regardless of the measures undertaken (Bordens & Abbott, 2018), prior to the intervention all participants were given a teacher-made writing pre-test. At the beginning of the intervention, the pre-test was administered to four groups with a total of 168 participants (control, experiment and two other groups of the same population). The aim underlies the introduction of the test for four groups was to examine the existence or not of the writing problems in the whole population and check the validity of the sample. However, for the experiment and control group the aim of the test is to spot the writing problems found.

The participants were asked to write descriptive and narrative paragraphs. The allocated time was one hour and half, estimating 45 minutes to be spent in writing one paragraph. This time length was intended to decrease in the coming sessions because one of the goals was getting students to write in due time. Following the IELTS examinations, 20 minutes is the due time predicted for testees to write 150-200 words, as a paragraph, to illustrate a diagram or a table (McCarter & Whitby, 2014). From 150 to 200 words is equivalent to a length of a paragraph because this latter was supposed to include from five to ten sentences (Arnaudet & Barret, 1981; Rumiseck & Zemach, 2003).

According to Knapp (2017), the post-test does not have to be the same as the pre-test. This implies that the same questions or themes do not have to be repeated. As a result, another teacher-created writing exam was used as a post-test to assess the effectiveness of the intervention. However, the genre was always respected. The written pieces were analyzed quantitatively through scoring and qualitatively through content analysis. When completing pre- and post-test analysis, the goal of the pre-test is to assess how the two groups were initially similar or different from one other and to detect whether there was a pre-test difference (Shadish, Cook, & Campbell, 2002).

2.4.2 Questionnaires

A questionnaire is viewed as a “form of interview on paper” (Kumar, 2018, p.191), which is used to collect information about people’s knowledge, beliefs, attitudes, and behavior (Gillham, 2008). It is also described as any written instrument that presents respondents with a sequence of questions or statements to which they are to answer either by writing out their replies or picking them from a list of available options (Mackey & Gass, 2005). The questionnaires used in this study contain mixed item types: open and closed-ended.

The rationale for using questionnaires in this study is justified by a number of reasons. First, following Genesee and Upshur (1996), the questionnaire is most useful when used periodically and systematically "before and after the treatment" because the information supplied is valuable for "planning and assessing whole courses and units" (p.127). This is what justifies the use of the pre- and post-study questionnaires in this study (see Sections 2.4.2.3, 2.4.2.4, and 2.4.2.5 below for further understanding). Second, questionnaires are supposed to be more reliable than interviews because they are anonymous and that encourage honesty in answers (Cohen et al., 2007). Third, questionnaires are less time-, energy, and money-consuming. Fourth, questionnaires help researchers to gain a broad view about the topic under study (Drever, 1995). Fifth, questionnaires can gather data from participants, especially when the information is not restricted to data production alone, to express themselves regarding their motivation, attitudes, beliefs, reaction to classroom instruction, and so on (Mackey & Gass, 2005). Pre- and post-study questionnaires were used in this study to collect appropriate data. They were designed based on the questions of the data gathering tools in the study by Lee and Coniam (2013); pre- and post-study teacher interviews and pre- and post-study student interviews. Before describing each questionnaire separately, their design, validity and reliability are tackled in the following Sections 2.4.2.1, 2.4.2.2, and 2.4.2.3.

2.4.2.1 Designing Questionnaires. To design a well-constructed questionnaire, the researcher should put himself in a "respondent's state of mind" (Bordens & Abott, 2018, p. xix). To avoid question wording (Nunan, 1992) and taking participants' level into consideration, the researcher tried as much as she could to formulate very simple and clear questions. To avoid obtaining answers artifacts for the elicitation tool itself (Nunan, 1992), leading and loaded questions have been avoided; neither confusing and complex questions nor double desired answers questions were asked. Following Boynton and Greenhalgh (2004) and Dörnyei and Taguchi (2010), a well-designed questionnaire contains multiple items in each area of investigation. In this regard, Stone (1993) suggests that a questionnaire should be "appropriate, intelligible, unambiguous, unbiased, capable of coping with all possible responses, satisfactorily coded, piloted, and ethical" (p.1264).

As for questionnaire items, Mackey and Gass (2005) believe that mixing both closed- and open-ended items in a questionnaire is preferable. The questionnaires of this study featured various mixed items formulated in different ways and taking different forms. To avoid boring the respondents and to elicit relevant data, the researcher avoided depending on only one type and form of questions. Dörnyei and Taguchi (2010), following their experience, see that only exceptional questionnaires have more than 4 pages long and require more than 30 min. to complete. This was the case for the post-study questionnaire administered to the experiment group. To recall, its length was due to the failure of the interview that was planned to take place instead.

2.4.2.2 Validity and Reliability of Questionnaires. Despite the fact that questionnaires are popular to collect, in most cases, quantifiable data, the construction of validity and reliability, however, "is highly a specialized business" (Nunan, 1992, p. 143). Reliability is defined as the ability of producing the same or highly similar results on repeated administrations and validity of questionnaire is measuring what it claims to measure (Boynton & Greenhalgh, 2004; Bordens & Abott, 2018). In accordance with this idea, Roger (1995, as cited in Bordens & Abott, 2018) asserted that if the results obtained after administering a questionnaire vary widely that means it is not reliable. On the other hand, Boynton and Greenhalgh (2004) see that using previously published and validated questionnaire save both time and resources, but there is no harm for the researchers to establish their own questionnaires. The same authors stipulated that using a published and a pre-piloted questionnaire does not mean it is either valid or reliable.

Regarding the questionnaires used in this study, it is of paramount importance to point to the following: (1) The researcher did not use just previously published questionnaires because this study is not a replication of previously published study (ies). Even if this was not the case, significant research was undertaken on previous and comparable questionnaires and interviews. (2) In this study, the researcher piloted both instruments to enhance validity of the questionnaires since piloting serves two purposes: "ironing out any design flaws that have been missed... and enabling a formal evaluation to be completed" (Stone, 1993).

2.4.2.3 Pre-study Questionnaire. The pre-study questionnaire was administered to profile the participants of both groups, 77 students, before embarking on the intervention. It focused on their beliefs and attitudes toward writing, their past experiences with writing classroom and writing classroom-based assessment in addition to assessing their levels in writing and their awareness about their strengths and weaknesses. This questionnaire was designed largely for the exploratory phase (Boynton & Greenhalgh, 2004; Hyland, 2009) to highlight attitudes and behaviors' concerns that may be followed-up later by more in-depth methods (Hyland, 2009). Following Genesee and Upshur (1996), this type of questionnaire collects information on participants' prior knowledge, previous experiences, and individual needs. As a result, the pre-study questionnaire was developed to meet the following goals:

- exploring the existence or not of problems in student participants' writing
- profiling participants' attitudes and beliefs in English writing and writing assessment
- exploring participants' awareness toward English writing
- exploring participants' preferences toward qualitative and quantitative assessment
- exploring participants' motivation to improve their EFL writing
- examining participants' beliefs and perception of English writing and writing assessment to plan the intervention
- exploring participants' awareness about the writing criteria

a.Description. The pre-study questionnaire is divided into five parts entitled consecutively: Background Information, Students' Experience with English Writing and Writing Assessment at Secondary Schools, Students' Attitudes toward English Writing, Students' Beliefs about English Writing and Writing Assessment, and Students' Needs and Awareness toward English Writing**.** The first section includes five questions about age, gender, stream in high school, the choice of studying English language, and their English mark in the BAC exam. The second section includes three questions addressing participants' experience at high school such as whether they practiced writing in the classroom and whether they were asked to assess what they had written. The third section consisted of six questions, such as whether they enjoy writing in English and what motivates them to do so. The fourth section consists of seven questions concerning their views, such as whether they believe to be skilled grammar users and whether they have large and rich vocabulary. Finally, six questions are included in the fifth section. They address issues, such as students' perceptions of feedback and their reactions to it, as well as their knowledge of the need to improve their writing (see Appendix.1 for further information about all of the items included in the pre-study questionnaire).

b.Pilot Phase. After developing the questionnaire, testing it out i.e. piloting it with a sample of a population (Creswell, Clrak, Guttaman, & Hanson, 2003) is pivotal to avoid problems with respondents' answers such as the difficulties to analyze the answers and collecting irrelevant answers or answers viewed without importance because of the respondents' misunderstanding (Nunan, 1992). To check its clarity, the questionnaire was piloted to eight students with varying levels of English to find any difficulties, namely wording, any item eliciting artifact data, or other problems. The questionnaire was piloted to students of various language levels, low, average, and good, to guarantee that it is properly understood by all participants when administered.

Before distributing the copies, the students were advised to keep silent. The goal was to avoid giving similar responses or assisting each other in solving any problems they encountered. Doing so might give the researcher an opportunity to collect as many questions as possible in order to refine the questionnaire. Participants were requested to (a) read the questionnaire thoroughly once before beginning to respond, and (b) ask all the questions they wanted to ask, regardless of their type or content. Various oral questions were asked during the piloting phase concerning the meaning of certain words as well as the formulation of some items. Enough time was allotted and all sheets were received in less than half an hour. Overall, only a few problems were spotted. During the piloting phase, the researcher went over all the responses provided and confirmed their relevance. She also collected all the questions asked during the piloting time. As a result, appropriate improvements and refinements were applied accordingly. This largely consisted of changing wording, and reformulating certain items and eliminating others.

c.Administering Phase. After piloting the questionnaire and making necessary modifications, the final version was administered to participants in the morning, around 8.45 a.m. to the control group and 10.00 a.m. to the experiment one. The same guidelines in the piloting phase were also followed in this phase (see Section b just above). After roughly 25 minutes, all copies were received. Despite the fact that participants were permitted to use Arabic or French if they were unable to write their answers in English, none of them did so, with the exception of one respondent, who used just two words in Arabic because he could not find their equivalent meaning in English.

2.4.2.4 Post-study Questionnaire for the Experiment Group. The research questions and the study background actually determine the type of questions to be asked in the questionnaire (Mackey & Gass, 2005). The post-study questionnaire was used to gather content data about the outcomes of the intervention regarding different aspects. It was divided into two parts because it was lengthy. Before opting for this questionnaire, an attempt to interview four participants was made but failed because the informants felt shy and could not respond freely. As an alternative, they preferred writing instead because they believed it would allow them to think freely and shift easily from one question to another.

a.Description. The post-study questionnaire was divided into two parts to avoid tiring the participants and to collect as much relevant data as possible. Some items were followed by further note-taking to delve deeper into some specific aspects of the intervention, such as explaining the scoring rubric. The questionnaire was organized into seven sections: Rubrics Presentation and Use, Self-assessment Performance, Anonymous Peer Assessment Performance, Classroom Assessment Interaction, Student Writing Progress, Gained Writing Features, and Feedback Delivery. The first part of the questionnaire included the first three sections while the second part contained the last four sections.

b.Pilot Phase. The post-study questionnaire was piloted in the same way as the pre-study questionnaire, following the same guidelines. To accomplish this, the post-questionnaire was piloted to two groups of eight students, each of whom completed a part of the questionnaire. It is worth noting that the group includes students with various language proficiency levels. The problems gathered were all about word meaning.

c.Administering Phase. The questionnaire was administered after correcting, refining and adjusting any items' problems for the sake of relevance, linguistic accuracy, and fluency. This phase was significantly longer than the one devoted to pre-study questionnaire. The participants were grouped into four sub-groups for easy management and subsequent further follow-up note-taking use. This latter was solicited to gather in-depth information on some questionnaire's items. In line with this idea, deMarrais and Lapan (2004) suggest to use interview and follow-up questions or probes to construct a complete picture about participants, based on the description provided.

To avoid problems with providing incomplete answers especially with lengthy questionnaires, Mackey and Gass (2005) see that the respondents should be given enough time and in case of limited literacy they could be offered the option of providing oral responses. In two different sessions with an interval of a week, the two parts of the post-study questionnaire were administered. Each sub-group, with an average of nine participants, received the questionnaire in a large classroom in Building C of the Department of English following the same guidelines as in the pre-study questionnaire. After the participant returned the questionnaire, a follow-up note-taking session was held with him/her. This has lasted from 3 to 7 minutes on average. The questionnaire was completed in 30 to 40 minutes by each participant. The total time spent completing the first part of the post-study questionnaire by all participants and doing the follow-up note-taking was roughly 4 hours for the first part and 3 hours for the second part. Working with sub-groups was highly beneficial to both the researcher and the participants since they had not been kept waiting in the classroom for too long.

2.4.2.5 Post-study Questionnaire for the Control Group. Because the control group did not undergo the AfL procedures, the post-study questionnaire directed to them may be seen as a subset of the post-study questionnaire administered to the intervention group. This means that it comprises only items pertaining to writing development and individual gains such as self-confidence and autonomy.

a.Description. This questionnaire contains nine mixed items, open and closed. The total number of questions is nine with sub-questions for the 1st, 3rd, 4th, 5th, and 9th question.

b.Administering Phase. This questionnaire was not piloted since the questions chosen had previously been piloted because it is part of the post-study questionnaire delivered to the experiment group. This questionnaire was introduced to the participants of the control group following the same guidelines as in the pre- and post-questionnaire destined to the intervention group. All the answers were delivered in less than 20 minutes.

2.4.3 Writing Rubrics

To stand as AfL, feedback information should be available and various actions based on feedback are at the core of effective learning (Black & Wiliam, 1998 a). The rubrics were designed to deliver constructive feedback to students involved in writing classroom assessment. To best understand the rubrics designed, the researcher gave detailed explanations about their design. We have to bear in mind that most research studies do not include extensive analyses of the design of the rubrics employed, and that most design features have not been systematically explored (Panadero & Jönsson, 2013). In her article where she reviewed studies of rubrics in higher education from 2005 to 2017, Brookhart (2018) found that there was no relationship between the type and quality of the rubric, and the studies' results and rubrics' design and content were not explained sufficiently. In any case, one has to retain that the rubric must be powerful enough to be able to catch the equilibrium of being both a simple instrument for marking but also thoroughly exhaustive to provide constructive feedback reflecting the learning goals (Cox et al., 2015).

2.4.3.1 Developing Rubrics for the Current Study. Designing rubrics, especially for assessing writing, is a painstaking process (Andrade, 2000). The effectiveness of a rubric in an educational setting is inextricably related to its development (Crawford, 2001). The rubrics used were instructional and analytic in nature, designed to involve students in the writing progress and development, and even co-created with them to align with the AfL core ideas and to react to many scholars who insist on including students in the design process (e.g. Hafner & Hafner, 2003; Andrade, 2005; Brookhart, 2013; Cox et al., 2019). The development of the rubrics for this study went through three primary phases: pre-, design, and post-design. Each stage is clearly discussed below.

a.Pre-design Phase. To construct rubrics, Stevens and Levi (2013) describe four key steps. First, selection gives an opportunity to know what objectives students should achieve, how to arrive at that, how students have responded to previous assessment and instruction, and why this assessment is created. The purpose of selection is to determine the type of rubric and whether to adapt an existing one or create a new one that is appropriate for the situation. The second step is listing the criteria and sub-criteria. This sheds light on the objectives to be accomplished. The third step is grouping and labelling. This step is based on the two previous ones in which similar performances have to be grouped to form the dimension of rubrics. The last step is called application, where the different parts of the rubric have to be assembled and combined. On the other hand, to assess rubrics' value, Turley and Gallagher (2008) suggest taking the following questions in account: (1) What is the tool for? (2) In what context is it used? (3) Who decides? (4) What ideological agenda drives those decisions?"(p. 87). Via asking these questions, the benefits, limitations, and shortcomings of rubrics would be highlighted. To design the rubrics of this study, the researcher had to make a decision upon the need of using rubrics in relation to participants' needs, making an extensive search in the literature to collect well-designed and developed rubrics, and finally deciding upon whether: to adopt a given rubric, to adapt, to create a rubric using existing rubrics, or to purely create a new one. All these steps had marked the pre-design phase.

First, to decide upon whether to use rubrics or not, two critical questions were asked: What motivates the researcher to use rubrics? And why is that? The participants were EFL first year students at University who had enrolled for a degree in English. It was the first time for them to be in a classroom where the language of instruction was English i.e. English-only Approach. They moved from Arabic-only approach in classrooms at high, middle, primary schools and English was taught as a subject in the curriculum. At university, students who specialize in studying English have various fundamental modules which are all taught in English. Among those modules, they have the writing module. They approach this module and have no idea what and how they will be taught. For example, they do not know what writing process is and what genres are. In the second year, they are supposed to be introduced to essay writing and in third year to academic writing. The short term and long term goals were clear for the teacher-researcher but not for the participants. The course goals were clearly outlined in the writing syllabus, and time was limited. As a result, the researcher decided to employ rubrics as specific assessment tools that reflect success writing criteria and make the purpose clear to them. According to Andrade and Du (2005), the purpose of rubrics is not limited to assessing learners but may also “teach as well as evaluate” (p.1). Moreover, Andrade (2006) sees that “rubric-referenced formative assessment can encourage mindful approaches to learning and free students from the tyranny of a classroom where the teacher is the sole judge of quality” (p.9).

Second, to spot participants’ writing problems to design appropriate rubrics that would help them promote their descriptive and narrative genres, they were asked to answer a pre-study questionnaire and write descriptive and narrative paragraphs. The first tool uncovered their beliefs, attitudes, and past writing practices and assessment helped to profile students while the second tool used to extract their real and concrete problems in writing. Third, conducting a thorough search in the literature to collect well-developed rubrics was another important step. To that end, the researcher focused on collecting some rubrics as samples to inspire the design. The rubrics used as an inspiration source were not all writing rubrics or all analytic instructional since having various samples can allow for a more expanded vision.

Finally, the reasonable recommendations ended up with are modifying, accepting a given rubric, establishing a rubric based on existing rubrics, or just constructing a new rubric. In this scenario, the researcher chose to establish rubrics based on past rubrics in order to best suit the circumstance under investigation. Below are the rubrics consulted and used as a reliable source to design the rubrics of the current study. Without describing them in a detailed way, the researcher, instead, prefers to hint at the major features, either positive, negative, or both, in each of them in addition to highlighting the major aspects that inspired the current design:

Diederich, French, and Carlton's (1961) Score-sheet. It is the first assessment scoring tool to emerge. It has detailed criteria, but the descriptors are not provided. The detailed criteria of this tool were the important feature that attracted the researcher.

Jacobs, Zinkgraf, Wormuth, Hartfiel, and Hughey's (1981) Rubric. Being the most widely used rubric in English as a second/foreign language writing (Janssen, Meier & Trace, 2015), this rubric is designed for scoring ESL composition. The researcher of this study sees that this rubric is holistic with unclear scoring strategy and qualitative assessment. This makes it difficult to be used by the participants. For example, in 'content' in 'excellent to very good' matching to score '30-27', some questions would be asked such as what one could comment between excellent and very good and what the exact score is. This rubric provided information about the description levels and the organization of the qualitative assessment in relation to the quantitative one. This rubric was highly relied upon as a source to design the rubrics of this study.

6+1 Trait® Writing Rubric. It is an online assessment tool that is updated periodically. Despite the details it contains, it seems overwhelming and impractical for the participants to use it as it is. Besides, it incorporates criteria like voice which may be unclear for them. This point has been discussed in the literature review chapter (see Section 1.6.1). In fact, many rubrics leave out the voice criteria because it is regarded to be too difficult to describe. According to Murray (2004), voice is the attribute that helps us to detect outstanding writing more than any other. The voice is a characteristic that learners should recognize by themselves because it cannot be taught.

It could only be assimilated via further practice and advanced composing writing analysis.

TOEFL Writing Band Descriptors and IELTS Dependent Writing Rubrics. Both rubrics are holistic destined for assessing international tests. The IELTS has well defined criteria while the TOEFL is less detailed. Instead of having five bands in the TOEFL rubric, the IELTS rubric includes nine. Consulting them gives an idea about how international experts see the various levels of writing and how they judge each one. This helped in aligning the rubrics to establish to the worldwide standards.

East's (2008) Scoring Rubric. It is similar to the IELTS rubric, but it has more criteria and fewer bands; eight bands instead of nine and five criteria rather than four. This rubric was created based on communicative competence. This aspect shed light on the possibility of incarnating rubrics in a given theoretical framework. It means that rubrics could also be principled in a set of tenets regarding the purpose and the objectives they are designed for.

Hibbard and Wagner's (2003) Analytic Narrative Rubric. The rubrics developed by these authors are so complex that students at pre-intermediate level cannot understand their meaning. They are destined for teachers because their language has a high register, but they can be reformulated to suit students' understanding level. This rubric provided details specific to the narrative genre which help in spotting the features specific to this genre.

Rubrics for Different Learning Outcomes. The nine rubrics used in the Office of Institutional Research and Assessment (2017) at the University of North Carolina are addressed for different learning outcomes. All rubrics contain a set of criteria with descriptors except three which contains a scoring strategy in addition. This shed light on pure qualitative rubrics and how to design rubrics destined for learning purposes.

The VALUE Rubrics. Valid Assessment of Learning in Undergraduate Education Rubrics are16 rubrics established by AAC&U (Association of American Colleges &Universities) addressed for different learning outcomes. They were put in 2009 and got higher attention by educators, teachers and researchers all over the world. They are good sources to consult before designing rubrics for teaching and learning because they are organized in different ways and and presented in different forms. This may inspire and help rubric developers.

To conclude, the pre-design phase was primordial to embark on the design process per se and motivated the researcher to challenge the rubrics already designed so that to develop rubrics of high standards. In accordance with this idea, Andrade (2005) sees that rubrics improve when being compared to published standards, or being commented by other colleagues. In agreement with this idea, the rubrics developed were feedbacked by the supervisor of this study.

b.Design Phase. The following measures were considered while developing rubrics for this study:

Understanding the definition of the rubric. Defining rubrics is crucial because that pinpoint the necessary dimensions and aspects to include. This has been fully discussed in chapter of literature review (see Section 1.6.1). The researcher opted for pure instructional analytic rubric at the beginning of the intervention and for the scoring one at the end to break participants' affinity to scores and scoring.

Making the difference between rubrics and other assessment tools. After defining rubrics, it is also important to know about other assessment tools such as rating scales, checklist and others to make a clear difference. In fact, there is a difference between rubric and other assessment tools, but the former is at the top of the pyramid. In agreement with that, it is concluded that "holistic scales, checklists, rating scales, and analytic scales can be used in rubrics" (Office of Institutional Research and Assessment, 2017).

Categorizing students' writing problems after having a clear idea about them. After gaining an understanding of participants' writing challenges through their written pieces and their responses to the pre-study questionnaire, rubrics were developed to treat student' writing shortcomings and focus on their strengths to assist them write successfully.

Deciding upon the components to include in rubrics. The rubric content is linked to the function it serves. The rubrics designed for this study were intended for learning purposes. In addition to the qualitative assessment rubrics, a scoring rubric was also used, albeit not at the same time with the first ones. The scoring strategy was dissociated and used separately as a scoring rubric. Thus, overall three main components are included: the evaluative criteria, quality definitions of those criteria and a scoring strategy (Popham, 1997; Andrade, 1997/2000; Reddy & Andrade, 2010). First, to decide about the criteria to include, the researcher asked the following questions: what are participants' weaknesses and strengths in writing? And what are the standards of good writing? It is of paramount importance to give "clear concise performance criteria and provide a forum in which students can create their own learning opportunities" (Cox et al., 2015, p. 27).

Second, the descriptions, or descriptors, and performance levels should convey the message appropriately; hence, paying attention to the content and the language is primordial. In line with this, O'Donnell et al. (2011) see that performance levels should be appropriate and understood to both teachers and learners. However, there is an issue with descriptions' number. According to Arter and Chappuis (2007), the number of levels is not yet fixed and it may range from four to six, or three to five (O'Donnell et al., 2011). Overall, from three to six levels are permitted; less than three may shallow the meaning of the rubric or cause it to lose its sense, and more than six levels could make the rubrics loaded and bothering, or even cause it to lose its function. Except for paragraph form, which is deemed a checklist item because it was regarded optional, the researcher chose four levels that were clearly understood and commonly known by the participants. Third, the scoring rubric was used to measure their performance and compare inter-raters correlation between self-, APA, and TA.

Deciding upon the rubric type. Collecting all participants' writing problems and deciding upon the purpose, which was AfL, determined the types of rubrics to use in this study. To help and motivate the participants improve their writing, analytic instructional genre specific rubrics were developed. This is not to discount the importance of holistic rubrics, but looking for details, improving weaknesses, and highlighting strengths in students' writing was best served by instructional analytic rubrics. According to Andrade (2000), instructional rubrics support student learning and development and act as standards-referenced assessment tools, making them ideal for classroom use (Brookhart, 2013) because they "help students understand what is wanted on an assignment, help students understand what a quality is... a product looks like, [and] enable students to self-assess" (Arter & Chappuis, 2007, p. 31), and help students to generate self- and peer feedback (Jonsson & Svingby, 2007). Moreover, in the EFL writing context, instructional rubrics are seen a way "to facilitate the writing process and effectively assess student writing" (Sundeen, 2014, p. 84).

Opting for an easy, clear, and understood rubric. To create an easy, clear, and intelligible rubric, the rubric itself, the rubric surroundings, and the beyond rubric should all be simple, clear, and understandable. This implies that students must be given clear and explicit expectations when embarking on expressive writing assignments (Sundeen, 2014). The rubric itself is the form and substance, but the rubric context is the aims for which it is used and why. Beyond rubrics, on the other hand, emphasizes that students' skills should not be limited to rubrics, but rather be viewed as a tool to help them promote and strengthen their abilities rather than being the sole way of improvement and progress.

Involving students whenever possible in developing rubrics. There are several advantages to include students in the construction and development of rubrics. Becker (2016), for example, reveals that incorporating ESL students in the creation and/or use of a rubric greatly enhances their summary writing ability. Involving students may aid in: (a) avoiding misunderstandings and misinterpretations prior to the proper application of the rubric; (b) increasing students' awareness of themselves as important actors in the educational process, triggering their abilities to recognize themselves and be more professional (Boud, 1991); and (c) reducing faculty workload by allowing students to do some of the work (O'Donnell et al., 2011). After explaining to the participants what a rubric is and its overall content, the researcher held an in-classroom face to face conference where together the teacher-researcher and participants discussed the rubrics' content and function and the overall classroom-based assessment. In accordance with this idea, Andrade (2000) and Payne (2003) see that the best source of feedback about assessment is to listen carefully to students. Moreover, Cox et al. (2015) find that rubric, unlike a marking template, is not only focusing on giving guidance to marker “but includes the student” (p. 26).

Caring about the validity and reliability of the tools. Validity and reliability is a core concern when designing both alternative and authentic assessment tools (Moskal & Leydens, 2000), including rubrics (Andrade, 2006). Validity and reliability of rubrics used in research have not received enough attention (Reddy & Andrade, 2010, p. 433), and few studies reported on the validity of rubrics (Rakedzon & Baram-Tsabari, 2017). The validity and reliability were not a main concern in their study since the scores generated by the scoring rubrics were not taken into consideration in the final mark, but these two concerns have been considered in this work regarding some dimensions like inter-raters correlation in self-, APA, and TA. To guarantee that the researcher went through numerous measures to increase the validity and reliability of the rubrics used for grading, the following criteria were given careful consideration:

1. using previous rubrics to design the rubrics used in the study. This maximizes the validity of the newly created tool.
2. formulating easy rubrics to handle with easy and clear language. Studies about the validity of rubrics have shown that language clarity and appropriateness is a central concern (Reddy &Andrade, 2010). This is to say adjusting the language to students’ level is very important.
3. piloting the tools different times to gather student feedback to improve the rubrics is very important.
4. discussing the rubrics with students whenever possible even while using them.
5. familiarizing and training students to using the rubric. Using a rubric without training may not improve the reliability or the validity of assessment (Razaei & Lovorn, 2010). In addition, the reliability and validity of rubric use is not only limited to teachers, but also it is the students’ business (Jonsson & Panadero, 2017) as they are at the end the proper users of it. Students must be shown how effective writing may seem through showing “modeling or written text” … “This

is an issue of fairness” (Spandel, 2006, p. 20). For example, showing models of good and poor writing was very helpful.

6. trying to use the rubrics with a large number of students. That was the case in this study, as the rubrics developed were used with 38 students.

Rubrics and time. The time necessary for rubrics’ usage should be neither shorter nor longer than the time it merits. These are key questions to ask for this purpose: Is the time frame appropriate for the rubrics developed? Is not using the rubric a waste of my and my students’ time? Paying attention to the time period is vital for two reasons: less time may not help in activating some rubrics, and using rubrics for an extended length of time might be uninteresting for students and hinder their development. For example, one of the drawbacks of the study by Andrade and Du (2003) was the short time length they used in the rubric; therefore, they ultimately advised a longer term. To deploy rubrics for an appropriate amount of time, the researcher relied on students’ responses when asked how many times they thought they may need to enhance their writing using the rubric before the intervention started (see Table 2.6 in Section 2.5.2.2).

c.Post-design Phase. Negotiating and discussing rubric’s design and usage with participants was primordial (Andrade, 2001) because it was, after all, destined to them, any problem might hinder their appropriate use and might affect their effectiveness (Brookhart, 2013). For further clarification, the writing criteria should be legitimized (Andrade & Du, 2007) to involve participants in further assessment performances, namely self- and anonymous/identified PA. Prior to the pilot phase, the participants were involved in the design, but at that time they had no background knowledge about this tool and little feedback was received. As a result, the researcher decided to proceed to introduce rubrics several times to the participants for further improvements. This aimed at: (a) having clear and in-depth definition of the tool, (b) stimulating participants’ serious thinking about the rubric, and (c) refining the model taking into account participants’ feedback. This phase served as a training stage where participants were familiarized with rubrics and student assessment practices.

After developing the rubrics, producing the final versions were subjected to pilot and discussion in-classroom. After that, they put into official practice. The following points were set to refine the final rubrics:

- exposing the final versions to students
- discussing them in classroom conferences
- exposing them to trials sessions where they try to assess their own and their peers' writing using the rubric
- collecting all the problems the participants faced while practicing using rubrics
- clarifying the short and the long term goals
- exposing the rubric after each refinement. This goes in agreement with Hafner and Hafner (2003), who argue that students should be involved in co-creating and developing rubric and even in modifying them.
- getting students' agreement to start officially rubric implementation

Rubrics were piloted three times. The problems found and the alternative solutions are displayed in Table 2.2.

Table 2.2 *Problems and Solutions for Rubric's Piloting*

Problem	Solution
Using descriptive adjectives like in the excellent use of vocabulary: the students can use a wide range of appropriate words.	Using the counting technique. For clear understanding see Appendix 4
The counting technique 1234	The counting technique 0123.
Memorizing the abbreviations and the descriptors	Planning a session in which both the researcher and the participants worked together. The participants were given an empty sheet to develop the rubrics content using their own techniques, such as using figures, arrows, and other symbols to memorize quickly the abbreviations and the descriptors. Then, this sheet was given to the participants on their demand

	each time they needed it.
Predicting problems (The researcher's sense)	Opting for a follow-up section associated with rubrics was used to track participants' problems, if any, while using rubrics and when performing assessment. In self-assessment, this follow-up section was a space for students to expose their problems in writing and their difficulties with using rubrics whereas in PA it was a forum of discussion where they expressed their opinion toward their classmates' writings and received their counterparts feedback in addition to criticizing rubric use in PA (see Appendix 5)

2.4.3.2 The Developed Rubrics. Because the rubrics were destined to classroom writing assessment, the entitlement at the top of the rubrics changed based on student assessment performance conducted; 'self-assessment/ assessor's full name' and 'APA/ code' (For further understanding, see Appendices 1 & 2). To rule out any confusion about the developed rubrics, the following questions help in further understanding them.

Why were abbreviations used instead of full written items? Abbreviations were used to accelerate students' memorizations of the items.

Why have you kept the descriptors separate from the rubric support? The descriptors were kept in a separate sheet to avoid producing lengthy rubrics, to make it practical to the user, and to help students memorize them as quickly as possible.

Why were paragraph form and title and handwriting kept separate? They are considered as technical aspects. By technicality, the researcher means that these elements represent the form and not the content of writing. This does not mean that they are unimportant. To clarify, indenting or not and caring about a block format is the appearance that the paragraph should have, but it does not seem to matter. That means this is not worth much practice to be assimilated and understood. As for handwriting, it is not that an important content point to be included as a mechanics criterion and be given a chance to contribute in the overall assessment of the criterion. If included in the mechanics of writing, like in Jacobs et al.'s (1981) rubric, having poor handwriting does not mean the mechanics are poor. In other words, if it is included, it will affect the overall assessment. It is true that if the reader cannot decipher the writer’s handwriting the content remains unknown and ambiguous, but the participants did not show dyslexic problems in writing and they had all passed BAC exam without being rejected because of their handwriting. So the purpose was to help them promote their handwriting and make it readable and, why not, nice without problematizing the issue.

For paragraph form being indented or not, both forms are correct. Deciding upon the former or the latter is up to the teacher, with/out participants, to agree upon which form to produce. In the department where the intervention took place, the writing staff, after holding a group discussion, agreed upon indenting the paragraph. Lastly, for paragraph title, the researcher had noticed that in the books about teaching writing to students, some paragraphs are entitled while others not. Given a title to a paragraph was a matter of choice. In a nutshell, these criteria can be omitted from the rubric without affecting its overall function, and this was the main reason which led to refer to them as technical aspects.

What is expected from participants to put in the comment section? The comment section, the boxes figured out under each criterion table, was kept up to the participants to judge their own and classmates’ works. It was not set with rules, but it could be done as such. As this had already been discussed with the participants, the researcher chose to leave it to their own choice. The researcher proposed to the participants that the lower qualitative assessment be posted in the comment box to motivate them to work more, but they refused. To explain the idea ‘comment follows lower in qualitative assessment’, for example in mechanics, if punctuation is excellent, capitalization is average, and spelling is poor, the comment is poor. Because they disagreed, the researcher decided to keep it open to them to freely comment because that will not affect the overall use of the rubric, and more importantly it was a chance to know about their reaction when given a chance to contribute into assessment. It was an initiative to make them feel partners in assessment.

What is expected from participants to put in the final assessment? Similarly and as previously discussed, the researcher proposed to participants to put low qualitative assessment in the comment section of each criterion and the same among the five criteria, but the participants disagreed. Accordingly, they were let to comment freely (The same as in comment section-the point just above)

Why was focus put on using the counting technique in formulating the descriptors? Using the counting technique was found easier to understand by participants than describing the descriptors. For example saying good vocabulary is ‘when students have a range of relevant words’. This was not understood. The participants were unable to assess using such descriptors. As a result, using the counting technique was considered as a more convenient way to make the content of the descriptors clear, simple and understandable. Furthermore, it helped participants in quickly recalling the descriptors and simplified the rubrics.

Why did you keep the scoring strategy separate from the rubric and use it at the end? The scoring strategy was kept separate as a scoring rubric and used at the end to break student participants’ affinity to scores as Algerian educational system is summative-driven praxis from primary to university level. The idea was to help students focus on their progress instead of wasting their time thinking about the scores and assuming that those scores represented their levels. To explore the importance of doing so, the participants were asked in the post-study questionnaire some questions followed by follow-up note-taking to get an in-depth understanding about that. For example, the participants were asked if adopting the scoring rubric right from the beginning would impact their qualitative assessment.

Which items are formulated out of the counting technique? Why? Some items were formulated out of the counting technique because of pre-determined objectives. This concerned knowledge of vocabulary, language constructions, and organization and content.

As for knowledge of vocabulary, the idea which lies behind this item was exploring participants’ richness in vocabulary. If we use the counting technique the same way as for the other items, it will be 0 new vocabulary, 1 new vocabulary, 2 new words, average, and 3 new words, and this is not reasonable. In fact, the same technique can be applied the other way round, but instead the researcher opted for 5 new words means excellent, 4 is good, 3 is average and 2 is poor to enhance students search for and use of more new words. This issue, on the other hand, was addressed in self-assessment as follows: excellent indicates that the participants used 5 new terms in their pieces of writing, 4 for good, 3 for average, and 2 or less for bad; however, in APA the numbers were maintained the same but were understood as vocabulary that the peer assessor could not identify. This indicates that if the peer assessor did not recognize 5 terms in the written work, the student writer was deemed to have an excellent vocabulary range. The same holds true for the other numbers.

In language constructions, the number of each sentence type was not limited for two reasons: (a) to put participants at ease while writing because when being constrained by a given number that could trouble them more because they were already struggling with writing; and (b) to raise their awareness toward using more complex structures and recognize that this can help in improving their language style. In general, organization and content aspects together with the items they included, were developed out of (most of them) the counting technique, because writing is a recursive and intricate process and limiting student thoughts and thinking as a whole would impede their advancement, interrupt their flow of thought, and disrupt their progress.

2.4.3.3 Writing Rubrics' Description. To describe the rubrics used in this study, the following elements are covered: rubrics' content and organization (see Appendix 2&3), descriptive and narrative genre-specific writing rubrics (see Appendices 2 &3), description levels/performance levels/ descriptors (see Appendix 4), and rubrics' follow-up sections (see Appendix 5).

a.Rubrics' Content and Organization. The rubrics are composed of two main components: rubric-feedback support and the follow-up section. The rubric-feedback support contains technical as well as content aspects. The technical aspects included in the rubrics are paragraph form, paragraph title and handwriting. These technical aspects are explained fully in section 2.4.3.2, just above.

Rubric content is presented into five tables entitled respectively: mechanics, vocabulary, language use, organization, and content. Each of these five criteria has items/attributions, and descriptors. First, for mechanics the attributions are spelling mistakes, capitalization and punctuation. Second, use and knowledge of vocabulary and foreign vocabulary are the attributions/items of vocabulary criterion. Third, for language use the items are constructions, subject-verb agreement, tenses, use of articles, use of pronouns, use of prepositions, use of negations, grammatical function of words, meaning at sentence levels, and sentence problems. Fourth, for organization the attributions are stated ideas, expressive ideas, connection between ideas, and space order or time order. Lastly, for content criterion, topic sentence, supporting sentences, concluding sentence, and relevance to the topic are the attributions. At the end of each criterion, a comment section was associated to give a chance to the participants to give their qualitative assessment. At the end of the five criteria, another final assessment comment was provided. This has been fully explained in Section 2.4.3.2, just above.

The rubrics were supplemented by a follow-up section which served in tracking participants' usage of the rubrics and also maintained discussion for further improvement and motivation. The follow-up section differed following student assessment conducted. In self-assessment, the participants were given a space in which they expressed what did and did not satisfy them in both using the rubrics and their writing, while in APA the follow-up section was addressed to both student writers and users of the rubric (anonymous assessors). The former were supposed to express what they thought about their peers' writing and whether the rubric did help them to assess and why (If not), or how (If yes). But, in APA this section was addressed to two persons: the student writer and the student assessor user of the rubric. This would help us to cross-check their inquiries in self-assessment through performing APA.

b.Descriptive and Narrative Genre-specific Writing Rubrics. The rubrics were designed and organized in accordance with the participants' involvement. First, the descriptive and narrative writing rubrics, as well as the scoring rubric, had the same format. The distinction was in genre-specific features such as space and time order, for example. In terms of structure and organization, the rubrics were divided into five tables for easy handle by participants. Each one corresponds to an aspect, or criterion, which are mechanics, vocabulary, language use, content, and organization. Each aspect included a set of items. For example, mechanics was formed of spelling mistakes, capitalization and punctuation items.

The organization and order of the criteria and their corresponding attributions or items was kept in the same order to avoid confusing the participants while applying them. For further clarification, in mechanics, for example, the order of items from left to right was kept the same in both rubrics which is spelling mistakes, then capitalization, and lastly punctuation marks. The same order was kept in the other criteria. Another important point to stress is the key characteristic features of each genre. They were highlighted to remind participants of their importance. The features were: tenses, space order or time order, connection between ideas, topic sentence, supporting sentences, and concluding sentence.

c.Description Levels/Performance Levels/ Descriptors. The description levels are mostly quantified for most items. This went in agreement with participants' understanding level because they could not perceive qualitative descriptions via using for example descriptive adjectives. To keep students involved in classroom-based assessment, the quantification was also discussed with them in light of their capacities and awareness. The number of mistakes was quantified, for most of the items, at different levels as follows: zero mistakes for excellent, one for good, two for average, three or more for poor. Some items were exceptionally treated as it had been explained just above in section 2.4.3.2.

d.Rubrics' Follow-up Sections. The rubrics, both in the descriptive and narrative genre, were associated with a follow-up section that differed according to the type of assessment conducted. In self assessment, because the focus was on the students themselves, the interest of the follow-up section was entitled 'Message from the **student writer** and **user of the rubric** to the teacher'. Within it, two sub-sections emerged 'My writing' and 'My use of the rubric'. The former included two blanks where students were given a chance to express their dis/satisfaction about their use of the rubric and their writing.

In APA performance, the follow-up section was different but the purpose was kept the same. In other words, it is maintaining the interaction student-student and student-teacher. To ensure anonymity, a space to write a code appeared at the top right of the sheet. The organization and the content as well were different. Instead of one section as it was the case in self-assessment, in this student assessment the same section was divided into two distinct subsections as follows: 'Message from the **student user of the rubric**' and 'Message from the **student writer**'. For further understanding, see Appendix 5.

2.4.3.4 Scoring Rubrics

a.Rationale. The scoring rubric was used after the instructional one to quantify the qualified performances because the main purpose was engaging students in classroom assessment to enhance their writing rather than focusing their attention on marking and marks. This goes in line with Black, Harrison, Lee, Marshall, and Wiliam (2004) who suggest that giving marks to students is a way "to compare themselves with others; those given only comments see it as helping them to improve. The latter group outperforms the former" (p. 18). Overusing scoring students' writings might be seen harmful than beneficial. However, the numerical outcomes can support the qualitative assessment if the purpose pre-set is using assessment to help students learn and progress.

Building on the literature review chapter, in order to synergize the effects of both quantitative and qualitative assessments, i.e. scores with qualities, a scoring rubric was introduced to help the participants quantify their writing performance and know about their quantitative progress; quantifying what had been qualified at the outset. The scoring rubric was introduced after exposing participants to qualitative assessment for a period of time. This was to minimize the participants' affinity to scores and instill in them the need for quality rather than quantity. Furthermore, this would teach them that grades are not the only method to assess their growth. Nearly all participants believed that the only way to determine their progress is scoring. To break that vicious circle of learning-grading adopted in their prior experience, the marks were distantiated and kept at the end.

b.Description. The scoring rubric was also analytic in both content and purpose. It was the same as the rubrics used for qualitative assessment but it included, in addition, a scoring strategy. It was analytic in content because it displayed detailed scores of the different items to sum up the whole for one final mark. It included intervallic scores. The intervals were marked continuously; for example attributing 1 for excellent, 0.75 for good, 0.5 for average and 0.25 for poor; the difference between the four levels is 0.25. Paragraph form, title, and handwriting were not marked as they were seen technical.

2.4.4 Writing Working Portfolio Project

As discussed in the literature review in Section 1.8, writing portfolio assessment has been used for the purposes highlighted at the outset. The main purpose for this research project was exploiting the portfolio for AfL using the working kind to help participants backtrack their works and develop ownership and self-responsibility toward their writing. This might help students take their weaknesses and strengths into consideration, trying to progress in the writing through working on the former and taking advantage of the latter to boost their self-confidence.

At the end of the intervention, the participants were asked to assess their written works, their use of the rubrics, and their self- and APA performance. To this end, a set of guidelines were set in the portfolio project brochure to guide the participants accomplish their mission. The brochure was divided into two major parts: the first one consisted of a table including four sections and the second part included a question asked to the participants to give their opinion about the portfolio project. The four main sections are students' writings, student self-assessment, student APA, and student re-writing.

2.4.5 Researcher Journal Report

This data collection tool was designed to enable the researcher to keep a close eye on student participants' time and how much was devoted to their writing and student assessment performance (self- and APA), as well as track individual inquiries about rubric use (descriptors, follow-up sections, and peer feedback) and overall classroom questions. The rationale for using this data gathering tool was to cross-check what the researcher had reported and observed with what students had written in the follow-up sections and post-study questionnaire.

The content of the journal report is divided into five sections. The first one is for participant writing time. It includes a table organized into four time intervals]0-15],]15-30],]30-45], and]45-60]. The number of practices is three for each genre: descriptive and narrative. Each time, the researcher wrote the number of participants delivering their writing in the corresponding interval. The second and third sections are devoted to time spent in self- and APA respectively. They are structured similarly to the first part. The fourth section, individual queries concerning writing rubric use, is divided into three sub-components: one about descriptor-related inquiries, one about follow-up sections, and one about offering peer criticism. Each sub-section contains a table that shows how many participants asked their queries. To recall, there are six practices in all, three for each genre under consideration. An area for writing the questions posed by participants in each practice has appeared in addition to the table. The fifth section about participants' individual inquiries about overall classroom assessment is arranged in the same way as the fourth section's sub-sections i.e. it comprises a table in which the researcher records the number of participants who have questions and inquiries, as well as an area where she write their questions (see Appendix 8).

2.5 The Study

The study evolved around preparing and teaching the content of the rubrics, preparing for the intervention, the intervention itself, and the post-intervention. These phases were planned in agreement with AfL. All of this was covered in one year (2018-2019); two semesters with an average of one three-hour session every week, including the reading skill, which was primarily intended to help writing because they are both literacy skills. As previously stated, the teacher of the writing course is also the researcher of the present study. This offered a vantage point of gaining an insider's understandings of participants' reaction to EFL writing (Hyland, 2009). Consequently, that helped in understanding how writing classroom-based assessment was approached by EFL students.

2.5.1 Pre-Intervention

The syllabus provided by the administration, the syllabus of the intervention and all the preparations prior the intervention are detailed below.

2.5.1.1 The Official Writing Syllabus. The syllabus taught by the teachers of writing at the department of English- University of Algiers 2 for the academic year 2018-2019 is presented below. The syllabus covered the course objectives, the content to be taught, and the teaching and assessment methods. The purpose behind presenting the syllabus is to discuss its content in relation to the study conducted.

a.Course Objectives. By the end of the year, students should be able to:

1. acquire the basic skills to express themselves in the written mode
2. distinguish between the process of writing and the final product
3. understand the different stages of the writing process
4. gain awareness of paragraph unity and coherence
5. distinguish between the different types of sentences in a paragraph
6. draw an appropriate outline and identify the exact order and organization of ideas in relation to the topic
7. write topic sentences that fit perfectly with the idea of the paragraph and its type
8. develop a bigger set of vocabulary needed in writing

b.Syllabus Content. The components of the writing syllabus are presented below in order from the first to the last session:

1. Introduction to the Building Blocks of Language within a Text: words, phrases, clauses, sentences
2. Types of sentences
3. Writing different types of sentences (simple, compound, complex, etc)
4. Coordination and subordination
5. Focus on language mechanics (punctuation and capitalization)

6. Pre-writing techniques (brainstorming, outlining, etc)
7. Introduction to the writing process: the during-writing techniques-drafting (The Narrative Paragraph)
8. The after-writing techniques – revision and edition, proof-reading
9. Writing outlines for narration and description paragraphs
10. Revision in writing for the first term exam
11. How to write descriptive paragraph
12. How to write narrative paragraph
13. How to write process paragraph
14. How to write a Comparison-Contrast paragraph
15. Polishing up style: Achieving parallel structure and avoiding wordiness
16. Opinion Paragraph
17. Revision in writing for the second term exam

The above themes were programmed for the writing skill. Bearing in mind that reading overlapped with the writing syllabus as they are both literacy skills, not all the classes of reading were related to writing, but some were purposively and directly linked to writing i.e. they were reading in form and writing in purpose. The components under consideration are:

1. Identifying different types of sentences within a paragraph (Topic sentence, supporting sentences, concluding sentence)
2. Using the context to understand the meaning of words
3. Recognizing text types and patterns of organization
4. Distinguishing the purpose of different types of texts

c.Teaching Method. In this section, the teaching staff listed the following two requirements while teaching writing:

- The “Pre-During-Post” framework (i.e. activities before the actual reading/writing; activities while reading/writing; activities after reading/writing)
- The use of different types of texts, classroom assignments and homework, and pair work and group work activities

d.Assessment Method. The assessment method, as the ‘writing staff’ suggested, explained just how to give the final mark in each semester. This latter included attendance, continuous assessment (mentioned as such), mid-term test, and end-of-term examination in semester.

Based on the official syllabus offered, the researcher has tried to analyze its content in relation to AfL. First and foremost, the syllabus was in line with the study's overall goals and an assessment method was included. Furthermore, a clear description of the objectives was presented and overall course content was suggested. However, the content of the assessment method was basically still flowing into being summative in purpose, despite featuring the term 'continuous assessment'. The syllabus was exempt from announcing directly and providing ways to deliver constructive feedback. This is to say that the qualitative assessment was not present anywhere and none of the quantitative assessment outcomes was discussed to be helping in improving students' writing.

2.5.1.2 Syllabus of the Intervention. At the outset, the objectives highlighted by the official writing module were established to make the intervention goes in line with them, precluding any possibility to prevent students from receiving their regular courses or meeting the objectives pre-set. With respect to that, the literal content was covered in agreement with the practices planned in the intervention underpinned by AfL procedures.

a.Major Content of the Intervention. The materials, activities and the overall plan of the intervention were prepared in advance. The overall content is the following:

1. taking the exploratory phase outcomes into account to embark on the experiment
2. designing the rubrics with participants' help
3. teaching rubrics' content
4. holding classroom discussions with students whenever necessary
5. familiarizing participants with the rubric and discussing it with them
6. performing self- and APA and discussing these performances with the participants
7. conducting the writing working portfolio project
8. collecting the outcomes of the intervention through post-study questionnaire

b.Activities and Materials Used. The teaching method, as made explicit by the teaching staff, was not constrained by a set of specific guidelines and let open for writing teachers' contributions; therefore, the researcher introduced assessment as a major component in teaching writing. The practices used were essentially rubrics, self- and APA to teach two genres: the descriptive and narrative. Keeping the syllabus content and objectives intact, the students were taught the content of the rubric that includes all the components of the syllabus using various activities, namely illustrations, modeling, and practices.

The activities used to teach the content of the rubrics were various: correcting mistakes, punctuating and capitalizing ready paragraphs, showing models of different quality paragraphs, asking participants to write different types of sentences and transform one type to another, asking them to write topic sentences through exposing a set of topics, teaching them relevance through exposing paragraphs with odd sentences, teaching them coherence and unity through practicing and providing them with appropriate connectors, exposing them to paragraph samples and highlighting its different parts, planning various activities about capitalization and punctuation marks, and correcting different problems in sentences such as fragments, run-on, choppy, stringy, sprawl, and lack of parallelism.

Table 2.3 below presents a brief summary of the activities and the materials used to teach the different descriptors to participants.

Table 2.3 *Activities and Materials Used in Teaching Rubrics*

Description level	Activity(ies)
Paragraph form	Showing samples of indented and unindented paragraphs
Handwriting	Showing scanned samples of different handwriting qualities.
Paragraph title	Showing example of excellent, good, average, and poor; Titles. Asking participants to write appropriate titles for ready paragraphs or for a set of topics.
Spelling mistakes	Using the dictionary;

	Exploiting texts of reading skill to make them aware of their mistakes.
Capitalization and punctuation	Punctuating and capitalizing ready paragraphs; Exploiting texts of reading skill to recapitulate briefly some rules related to punctuation and capitalization.
Use of vocabulary	Meaning in context. This was included in the reading courses where participants should infer and recognize the meaning of words in context through using a set of activities.
Constructions	Recognizing the different types of sentences; Making the difference between the different constructions in reading sessions; Completing incomplete sentences; Writing different types of sentence about a set of topics.
Subject-verb agreement Use of articles Use of pronouns Use of prepositions Negations	Presenting sentences and paragraphs to correct these types of mistakes.
Grammatical function of words	Eliciting different words with their grammatical functions in texts. Meaning in context.
Meaning at sentence level	Checking the writing style in sentences, both correct and incorrect; Examining sentences written by students and analyzing their content in relation to their meaning.
Sentence problems	Checking the different sentences problems either in spare activities or in their own writings.
Stated Ideas	Showing paragraphs with undeveloped ideas and asking them to develop them.
Expressive Ideas	Making the difference between powerful and weak ideas in paragraphs and texts in reading sessions
Connection between idea	Teaching the connectors; Explaining their functions in texts
Space order/order of	Teaching students specific transitions of space, order of

importance and time order	importance, and time order. Presenting paragraphs with missing and inappropriate transitions to fill and correct.
Topic sentence	Highlighting the topic sentence for a set of paragraphs; Asking them to write topic sentences for a set of topics.
Supporting details	Presenting paragraphs to highlight their supporting details Presenting paragraphs with missing details to complete
Concluding sentence	Presenting paragraphs to highlight the concluding sentence; Presenting paragraphs with missing concluding sentence and asking them to complete them; Presenting a set of topics to develop for them a concluding sentence
Relevance to the topic	Presenting paragraphs with irrelevant sentences to cross out.

The main difference with the control group is that the syllabus was transferred as it was provided by the writing staff. It was taught following the same guidelines included in the brochure offered by the writing staff, without using or presenting any rubric content and without implementing any form of AfL. In addition, the teaching method and the corresponding activities set by the staff were respected carefully and the objectives were also strictly taken into consideration.

2.5.2 Preparing for the Intervention

Planning for the intervention proper began once the rubrics' content was taught. The major concern was discussing the findings of the rubric pilot phase and the general setup of the classroom-based assessment environment.

2.5.2.1 Discussing Rubrics' Content. The outcomes of the pilot phase of rubrics were discussed in the classroom orally with the help of the whiteboard. The discussion concerned the quantification or the counting technique as it was the main idea of the rubrics. For recall, the main idea or the basis for establishing most of the description levels was based on concretizing them through quantifying mistakes. The quantification was referred to as 'the counting technique' as it was also called by students the ‘0123-technique’. It means: Excellent: zero (0) mistake, Good: one (1) mistake, Average: Two (2) mistakes, Poor: three (3) or more mistakes.

The counting technique was carefully discussed with the participants. At the beginning two suggestions, ‘0123-dechnique’ and ‘1234-technique’, were exposed to participants to test their awareness and the extent to which they would allow themselves to make mistakes. The participants were given the opportunity to vote on which technique to use four times, and the results are shown in Table 2.4.

Table 2.4 *Participants’ Choice for the Appropriate Technique*

The technique **Vote number**	**‘0123-technique’**	**‘1234-technique’**
1st time	66.67%	33.33%
2nd time	43.33%	56.67%
3rd time	53.33%	46.67%
4th time	60%	40%

The participants were mostly in favor of the first option as they were aware of the importance of getting rid of their mistakes. For long-term use, a few participants noticed that this technique could be further developed, especially when they advance in their writing. Most of them showed motivation to progress and write better pieces. For example, one student said, ‘how come we tolerate ourselves to do lot of mistakes. We are learning to get rid of mistakes and not to do more mistakes’. Another student said, ‘even though that seems hard, we have to be strict with our writing’.

For the second group, supporters of 1234 technique, few participants’ fear was how to get rid of mistakes drastically right from the beginning because they were tolerant with themselves toward making mistakes. In other words, they were quite convinced that expelling mistakes should take time. They justified their stance as being beginners and not doing mistake was challenging but rather legitimate. For example, one student said, ‘it is impossible to limit ourselves to do few mistakes’. Another student said, ‘I have problems with spelling mistakes. I do a lot of them. I am wondering how to apply this technique’.

After the in-classroom conferencing discussion, the participants were polled on their perceptions of failure to decide upon which technique to use. Six suggestions were written on the whiteboard to vote on. These were the results:

1. I disappoint easily, and I give up (6.05%)
2. I disappoint easily, but I don't give up (44.45%)
3. I don't disappoint easily, but I give up (13.89%)
4. I don't disappoint easily, and I don't give up (16.67%)
5. I don't disappoint at all, but I give up (6.05%)
6. I don't disappoint at all, and I don't give up (12.89%)

Together, teacher-researcher and participants, decided to adopt the first option '0123 technique' as the majority of the student participants do not give up if they make mistakes. To confirm this choice in the familiarization step, the participants were given the chance to apply either the first or the second suggestion and most of them (83%) challenged and adopted the first proposition, 0123 technique. The discussion was conducted for the following reasons:

1. making participants feel important in the classroom. This is through contributing in decision making as long as "the classroom assessment environment" is seen as a "sociocultural reality experienced and interpreted by individuals", and the instructional entity is part of the experience including "interpersonal relationships and internal thoughts and feelings" (Brookhart, 1997, p. 162).
2. helping them face their responsibility toward their writing
3. divulging the importance of making mistakes but not abusing in doing so
4. making them aware of the long term goals and correcting their vision toward seeing just the short term ones
5. giving them another chance to understand the rubrics properly
6. having an idea about their perceptions of the rubrics before the proper use
7. maximizing validity and reliability of the scoring rubric

2.5.2.2 Writing classroom-based Assessment Management. As participants were supposed to be at the core of classroom in launching discussions and being involved in decision-making, a written questionnaire and face to face oral conferences were introduced to organize the writing classroom in order to plan for the intervention.

a.Written Questionnaire. In line with the AfL underpinnings, involving participants to decide about their progress in writing was deemed pivotal to help them at the end take charge of their own progress. Toward that purpose, the participants were involved to decide upon the following points regarding the overall intervention: (a) the choice of anonymous or identified PA, (b) the way of ordering the performances, self- and PA, within a genre and how to order performing genres, and (c) the duration required to master the use of rubrics. To do this, a written questionnaire was introduced associated with classroom conference to elicit necessary information for classroom assessment management. This discussion was conducted due to the complexity of the intervention. Following their answers in the pre-study questionnaire, the intervention was complex regarding participants' level as being first year EFL students and being unfamiliar with self-assessment, PA, rubrics, and overall AfL. The results of the questionnaire regarding the points pre-set above were: First, in doing PA, all students chose anonymity, i.e. APA, for the following reasons: (a) avoiding social conflict and protecting their friendship, (b) avoiding subjectivity in assessment, and (c) avoiding carelessness when performing identified PA. In relation to anonymity, the participants were also asked whether they wanted to know the assessor or not and why. The answer to that is displayed in Table 2.5 below.

Table 2.5 *Post-anonymous Peer Assessment Suggestions*

Items	%
Yes, after each assessment I want to see who assessed my writing.	28.95%
Yes, but after we finish all the PA sessions.	18.42%
Even after we finish all the PA sessions, I don't want to know who assessed my writing.	52.63%

The first choice was justified as follows: (a) checking my grammar mistakes, (b) discussing my mistakes with the assessor, (c) satisfying my curiosity, (d) seeking further advice, (e) asking about the clarity of my handwriting, and (f) evaluating my progress. For the second choice, the arguments were: (a) checking whether the assessor’s comments worth to be taken into account or not, (b) satisfying my curiosity, and (c) knowing the reasons behind assessing my writing as such. For the third choice, the arguments were: (a) caring about assessment for the sake of correcting mistakes more than the assessors themselves, (b) avoiding social conflict, and (c) having no curiosity for that. At the end of the questionnaire, an open question for additional information was added, but just one participant answered saying: ‘I think that self- and PA will be so interesting. They are new techniques that would give me a chance to improve my writing.’

Second, when asked about the possibility of mixing both performances self- and PA or performing each for a given period before moving on to the second one, nearly all participants wanted to mix the two performances, self- and APA, for each genre but not the two genres for two main reasons: (a) avoiding boredom and breaking the routine, and (b) benefiting from their peers’ feedback and advice to reflect on their self-assessment and their writing in general. They suggested beginning with self-assessment and then APA to compare their assessment with the one of their classmates to avoid being impacted with APA if it is received the first.

Third, when asked about the possibility of mixing both genres or not, the majority agreed upon not to mix the two genres for the two following reasons: (1) the participants were afraid of mixing knowledge about the two genres and that would slow their progress, and (2) they felt secure when they work in an organized manner.

Fourth and last, to enquire exactly about the number of times they thought they wanted to perform self-assessment, APA, and using rubrics, Table 2.6 displays the means of the different performances according to student participants’ needs.

Table 2.6 *Number of Times Required for Using Rubrics and Performing Self-and Anonymous Peer Assessment*

Performance	**Self-assessment**		**Peer-Assessment**		**Rubric Use**	
Genre	Descriptive	Narrative	Descriptive	Narrative	Descriptive	Narrative
Means (Times)	3.05	3.52	3.08	3.16	3.45	3.77

The overall Means	3.34	

b.In-classroom Conferencing. The in-classroom conference was conducted for multiple reasons: (a) clarifying the ambiguities related to assessment performances to, (b) clarifying the overall purpose of classroom assessment, (c) discussing mis/understandings, if any, of the goals, (d) giving participants another chance to expose their problems regarding rubric content and design and its interest, and (e) providing supportive advice and encouragements to relieve their stress toward studying writing.

2.5.2.3 Familiarizing Participants with Rubrics. After discussing rubrics' content and design, these assessment tools were introduced to participants to familiarize them with their usage and refine them, if any adjustment were needed to apply. They were tried two times using both genres. Each trial lasted more than 45 min. The participants found them difficult and hard. An oral discussion was held with students while performing the assessment using the rubrics, and guidelines and further clarifications were thoroughly provided when necessary.

2.5.3 The Intervention Proper

The intervention took place after preparing for it by designing rubrics, familiarizing students with them, and managing the classroom-based assessment involving the feedback. The emphasis was on two genres, descriptive and narrative, and students were asked to provide criticism on their own and their peers' works to assist them progress. Each genre was practiced three times with three self-assessments and three APAs.

2.5.3.1 Pre-requisite Guidelines. A set of three pre-requisite guidelines were strictly followed for the careful implementation:

- giving student participants enough time to assess using the rubrics and recording the time needed for that
- banning students to speak with each other and keeping enough distance between them to direct their questions to the researcher

- avoiding writing and self-assessing the same paragraph in the same session and postponing it for the next session. The aim was to give them time to think about what they had written.

2.5.3.2 Assessment Performances. As discussed in the literature review chapter, self-, APA, and TA were integral parts in the AfL used in this study. The two first performances were introduced a main performances in the intervention while the third practice was served to be scaffolding.

a.Teacher Assessment. TA served to scaffold classroom assessment. This included tailoring rubrics' usage to participants' understanding, clarifying how to deliver and respond to feedback, and assisting in student assessment via answering all participants' inquiries. In addition to written feedback, oral feedback was highly emphasized to adjust student assessment performances. Guided writing using rubrics had been also conducted in the classroom. TA served to maintain communication with participants to foster them expose their problems and bridge the gap that is believed to be weak and not fruitful.

b.Self-assessment. In self-assessment, the participants were given the written pieces, they had already written, to assess using the rubrics developed. They were given enough time to be able to do their work and to feedback in the follow-up section carefully.

c.Anonymous Peer Assessment. When compared to self-assessment, performing APA was more difficult. To execute it well, three elements were taken into account when developing APA:

- observing and examining the relationship between participants and their seating positions to be able to pair them appropriately. The participants who constantly seated next to each other were not partnered in APA to discard any doubt about recognizing their classmate's handwriting.
- dispatching the participants and asking them not to speak with each other

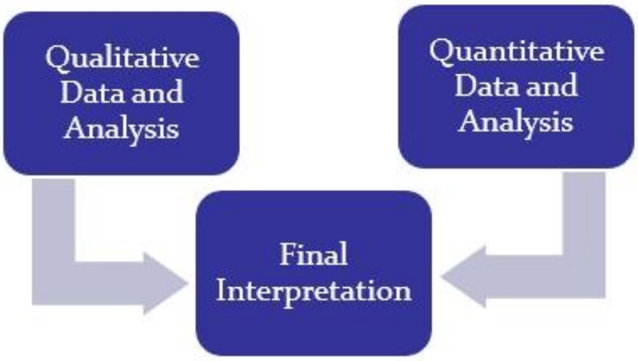
- allowing participants to give their own further remarks to their peers on their willingness after finishing PA and providing the feedback in the follow-up section

Having a classroom with mixed English levels was what complicated PA performance, and that was one of the major reason for opting for anonymity. To perform PA fairly, the researcher opted for combining one participant with high level (the assessor) with low level (the assessee) in the first practice, high with high and low with low level in the second practice, and random assignments in the third practice.

2.6 Data Gathering Procedures

To gather relevant data from the intervention conducted, three main phases were identified: pre-intervention, intervention and post-intervention. Taking into account the typologies proposed by Creswell, Clark, Gutmann, and Hanson (2003), the four factors as discussed by Creswell (2009) were taken into account. First, regarding timing, both data, quantitative and qualitative, are collected concurrently, i.e. at the same time because they are equally important in this study. A second factor that has been pointed out to by the researchers, Creswell et al. (2003), is weighting. This latter refers to the prominence conferred to quantitative, qualitative or both types of data. The study objectives, presented in section 2.1, and the nature of the dependent variable at hand need to be visualized equally. As AfL is a construct that is maintained by both formative, summative and the common zone between the two sides (Allal, 2016), this requires the need to both qualitative and quantitative assessment. A third aspect that has been evoked is mixing. It is defined by two questions: "*When* does a researcher mix in a mixed methods study? And *how* does mixing occur?" (Creswell, 2009. *Author's italics*). The answer to the first question is that mixing the methods happened in collecting, analysing and interpreting the data while for the second one the way of mixing them appears in triangulating and establishing validity and reliability of the research. The fourth aspect is theorising or transforming perspective. The purpose behind this is seeking whether the entire research is guided by a theory. This study is contextualized in theories such as socio-constructivism, autonomy and motivation. These are suggestions without being clearly and further explained how that is bordered in a given theoretical frame. The concurrent triangulation design was adapted to collect data without weighing which data is bigger as both are considered of equal in importance.

Figure 2.1 *Concurrent Triangulation Design. From Creswell (2014)*



The pre-intervention phase's main sources of data were the pre-study questionnaire, the pre-test, and the content analysis of participants' writings, whereas the intervention phase's main sources of data were rubrics associated with follow-up sections, as well as self- and APA performances, and the researcher journal report. The post-study questionnaire, post-test, and working portfolio were the primary data eliciting tools throughout the post-intervention phase.

2.7 Data Analysis Methods and Procedures

Both descriptive and inferential statistics were used to analyze the data collected for this study. However, one should bear in mind that statistics are not perfect; nonetheless, it helps researchers reduce uncertainty to better understand what they study (Knapp, 2017). The same author emphasizes that the statistical results address the whole group studied, not just individuals, saying "suppose we find that the average age within a group is 25: this does not mean that we can just point to any one person in that group and confidently proclaim "you are 25 years old" (p. 58). For this reason, it is critical for a researcher to know in advance what analysis procedures should be followed before collecting the proper data (Nunan, 1992). In this study, statistics are not used to generalize the findings obtained, and the results are destined to answer and discuss the research questions within the limitations of this study (For further understanding, please consult the introduction of Section 4.1 in Discussion Chapter). Furthermore, inferential statistics are also used to include a confidence level in the results (Sullivan, 2018).

Understanding research and the type of data available have an important role in determining which statistical test to use (Knapp, 2017). The data of this study were treated differently; manually, using Excel software, and SPSS (Statistical Package for the Social Sciences) version 26. Manual analysis was used for content and thematic analysis, which had proceeded through reading, rereading, and analyzing the texts obtained from the open-ended items in the pre- and post-study questionnaires, the second part of the portfolio, and the researcher's journal report. Table 2.7 shows which analysis procedure was used for which data collection instrument.

Table 2.7 *Which Analysis Procedure for Which Data Gathering Tool*

Analysis method	Which data gathering tool to be analyzed
Manual analysis	Used for establishing categories and themes after analyzing the corpus of data obtained from the open items of the pre- and post-study questionnaires, portfolio project, and researcher' journal report, and the written pre-test.
Excel software	Used for running analysis of the closed-ended items of the pre- and post-study questionnaires, portfolio project, and researcher' journal report.
SPSS version 26	Used for Pre- and post-test scores.

In what follows, the researcher justifies the methods used without delving deeper into them so as to avoid creating other research topics. In this regard, the next part provides brief conceptualizations of the various methods used and categories as well as themes' nomenclature and levels relied.

The major analysis methods used for open-ended items were content, thematic, and blended (Also known as mixed and hybrid) method. As for content analysis, Weber (1990) states that the central idea is that "the many words of the text are classified into much fewer categories" (p. 12). This is to say that the main idea is to have fewer categories without causing any harm to the data. This require, in some cases, to put unclassifiable elements as they are, i.e. as distinct categories (Gillham, 2008). For thematic analysis, Joffe (2012) has reported that the good thematic analysis must "describe the bulk of the data" (p. 18) and reflect the contextual meaning. Another hotly debatable topic in the literature is the distinction between the two methods. This confusion has been basically rooted in history. The first method, content analysis, is basically quantitative (Smith, 2000) and the second one, thematic analysis, is qualitative (Neuendorf, 2019).

As far as naming the categories or themes in content and thematic analysis are concerned, it is of paramount importance to pinpoint that scholars found difficulties to draw a clear distinction between the two terms. To rephrase it, there is no clear difference regarding which word is best describing either method. Some scholars use theme in thematic analysis (Joffee, 2012), while others use it also for content analysis (Neuendorff, 2019). Others use both words interchangeably (Morse, 2008). For further understanding, a theme tends to be more abstract and implicit which needs further explanations (Sandelowski & Barosso, 2003/2007) whereas a category provides explanations of the text content and describes the participants' account (Gray & Densen, 1998). In this study, the researcher used the terms 'category' for content analysis and 'theme' for thematic analysis. When 'theme' and 'category' emerge simultaneously, this indicates that the method used is mixed: content and thematic analysis.

Another point of contention is whether the category or theme's depth is latent or manifest, also known as semantic. Braun and Clarke (2006) distinguish between top-down or theoretical analysis, which is driven by specific research question(s) and/or the analyst's emphasis, and bottom-up or inductive analysis, which is driven by the data itself. In this vein, the two authors have made a distinction between the two levels as follows: semantic level "…within the explicit or surface meanings of the data and the analyst is not looking for anything beyond what a participant has said or what has been written"(p. 84). In contrast, the latent level looks beyond what has been said and "…starts to identify or examine the underlying ideas, assumptions, and conceptualisations – and ideologies - that are theorised as shaping or informing the semantic content of the data" (p. 84). Furthermore, Vaismoradi, Jones, Turunen, and

Snelgrove (2016) have discussed the level of the theme and category, claiming that in thematic analysis, researchers consider themes with latent meaning whereas categories with manifest meaning while in content analysis researchers select between them before embarking on higher analysis levels. In this study both types are employed. It is top-down because the research questions are important, and bottom-up because the data obtained is equally important as the research field of AfL and writing having few studies. More precisely, as for the inductive kind Joffe and Yardley (2004) argued that it is primarily employed in new fields of study, despite the fact that it is not entirely inductive or data-driven. Another point worth emphasis is that the themes and categories generated from the data are tightly related to the scope of the study and making the distinction between level of the category or theme in this study has no impact because the major aim is to explore, investigate and report all the details, either major or minor.

To sum up, the content, thematic, and blended analysis methods used in this study, are displayed in Table 2.8.

Table 2.8 *Content, Thematic and Blended Analysis Methods Used*

Analysis method used	**Unit name**	**Unit level**	**Analysis method predominance**
Content	Category	Manifest/Latent Inductive (data-driven)-deductive (theory and research questions-driven)	Basically quantitative
Thematic	Theme	Latent/Manifest Inductive-Deductive	Basically qualitative
Blended method	Both	Manifest-latent Inductive-Deductive	Both

2.7 Chapter Summary and Conclusion

This chapter is dedicated to design the intervention conducted for the purpose of examining the effectiveness of AfL in relation to students' writing progress via implementing rubrics and student assessments, self- and PA, under teacher guidance and assistance. Toward this aim, a quasi-experimental design based on mixed methods was adopted to devise the study in light of the objectives and the research questions formulated at the outset. And, a non-equivalent comparison group design, including two groups of 77 participants of first year students from the department of English in the University of Algiers 2, served as the study sample. Furthermore, a concurrent triangulation design was chosen to collect the data of this study, addressing both qualitative and quantitative data as having equal importance. Pre- and post-tests, rubrics with follow-up sections, pre- and post-study questionnaires, a researcher journal report, and a working portfolio were used to collect data. These tools were used to elicit information for the study, taking into account all aspects of the intervention. The tools were carefully designed and piloted as necessary to confer due credibility for the findings and cover complementarity and triangulation, as being the utmost goal sought for via using mixed methods. After data collection stage, the next chapter is reserved to presenting and interpreting the results yielded by each data collection instrument.

3

CHAPTER THREE: DATA ANALYSIS AND PRESENTATION OF RESULTS

After blueprinting the study occurrence, explaining what design should be adopted, and why and how that should be applied, this chapter presents the analysis of the data gathered for this study. The data analysis process was divided into three main stages. In the first stage, the analysis process took into account the pre-study questionnaire and written pre-test, whereas in the second stage the data gathering tools concerned were rubrics with their follow-up sections, including self- and APA associated with peer feedback, and the researcher journal report. The last stage of the analysis tackled the pre- and post-test, post-study questionnaire, and working portfolio.

3.1 Data Analysis and Presentation: First Stage

The data analyzed in this stage were obtained from pre-study-questionnaire and written pre-test. The analysis methods used are presented in Appendix 14.a.

3.1.1 Pre-study Questionnaire Analysis

The pre-study questionnaire, for recall, was a mixture of closed and open items, organized into five sections. It was used to profile participants, targeting their attitudes, needs and awareness toward writing, and exploring their past experience with writing assessment. Concerning the open items, their answers were analyzed using content and thematic analysis methods (Discussed in Chapter 2, see Section 2.7). The responses of the closed items were reflected through percentages and numbers; numbers were used with closed items and percentages with open-ended items because the participants displayed more than one answer. To clarify, ‘I don’t know’ option that appeared with ‘Yes/No’ in closed-ended items meant that the participant could not decide whether it was ‘Yes’ or ‘No’ i.e. undecided. This option and ‘a little’ option had given an opportunity to participants to express themselves instead of forcing them to say either ‘Yes’ or ‘No’. The figures obtained through using Microsoft Excel were selected in terms of clarity and explicitness. In other words, any figure type, pie chart, histograms, etc., was not preferred over any other.

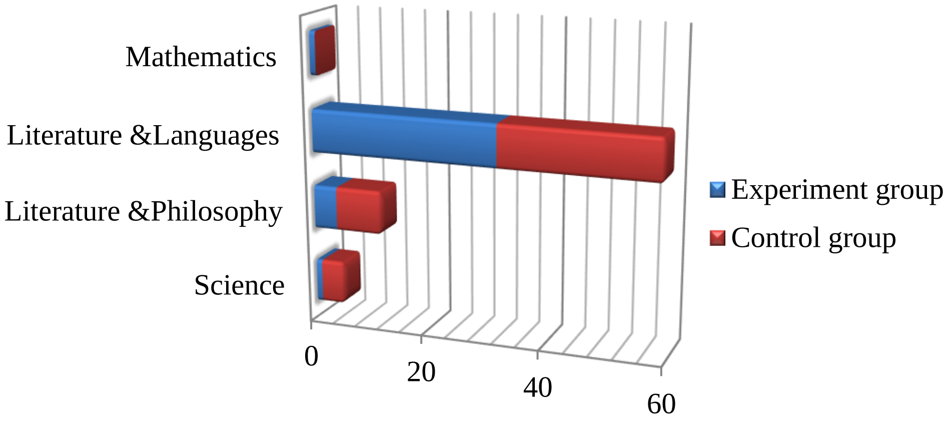
3.1.1.1 Analysis of the First Section: Background Information

Question 1: The mean age of the participants was calculated by summing all of the ages and dividing the result by the total number of participants. The mean age of the experiment group was 18.7, whereas the control group was 18.2.

Question 2. The number of males was higher in the experiment group than in the control one; 14 to 6 male participants respectively. Gender difference was not taken into account in this study, but reference to it appears in the section of recommendations.

Question 3. The majority of the participants studied literature and languages in high schools. None of the participants in both groups came from mathematics stream. Figure 3.1depicts the results

Figure 3.1 *Participants' Stream in High School*



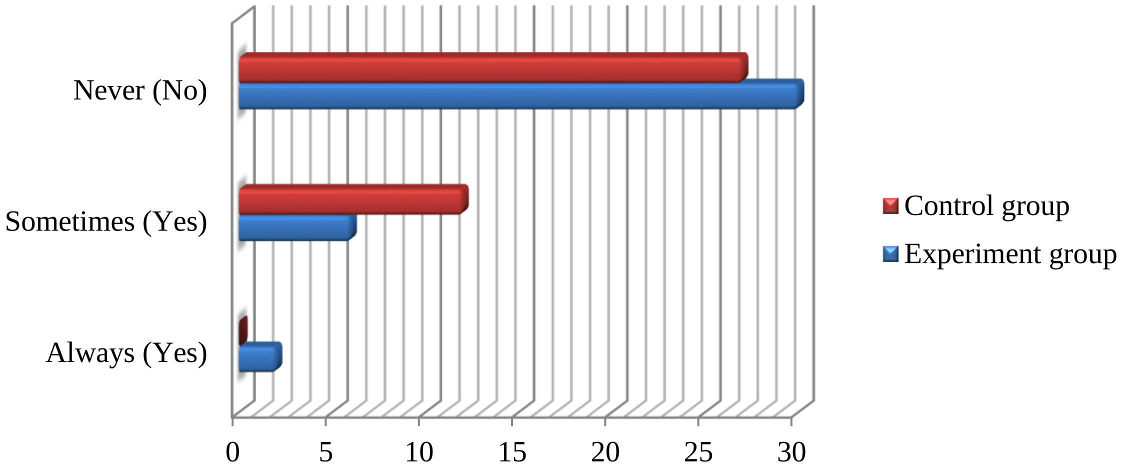
Question 4. Except for two in the experiment group and one in the control group, all participants in both groups freely chose to study English as a specialization at university.

Question 5. Grade mean of English language subject in the Baccalaureate exam was 14.12 in the experiment group, excluding two students who did not mention their marks, and 14.78 in the control group.

3.1.1.2 Analysis of the Second Section: Experience with English Writing and Writing Assessment at Secondary Schools

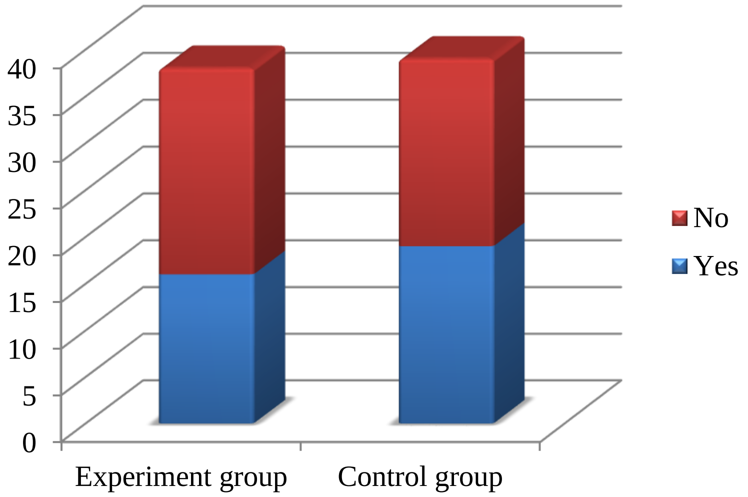
Question 1. Nearly, all the participants said that they did not practice writing during their regular classes in high school and just few of them responded positively. And only two of them in the experiment groups responded that they did it on a regular basis.

Figure 3.2 *Writing during Classes in High School*



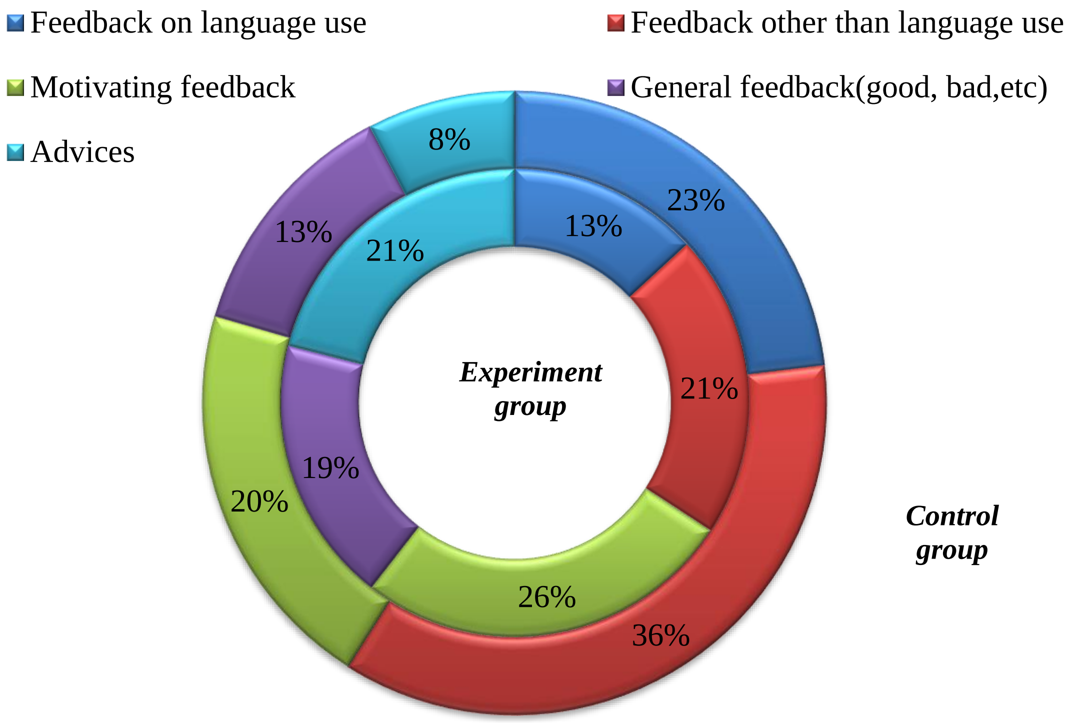
Question 1.1. The participants who responded ‘Yes’ when asked about whether their high school teachers provided them with evaluation were higher in the control group than in the experiment one.

Figure 3.3 *Presence of Writing Evaluation in High School*



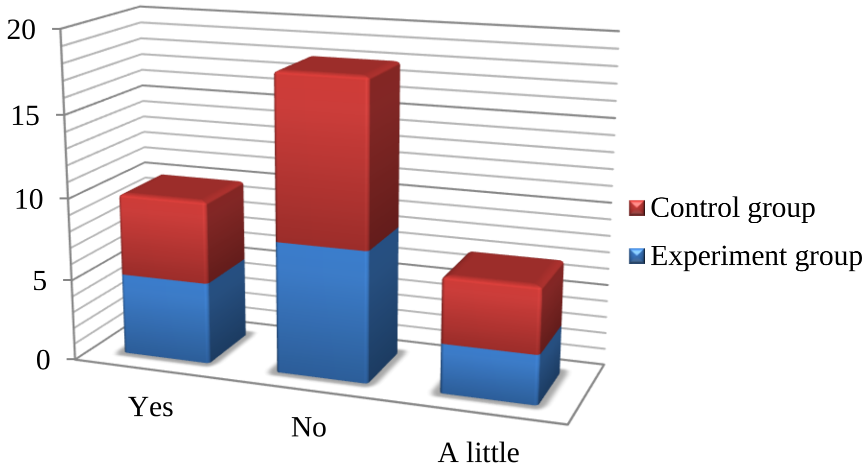
Feedback delivered is classified as 'feedback on language use', 'feedback other than language use', 'motivating feedback', 'general feedback', and 'advice feedback'. Further details are displayed in Figure 3.4. The outside circle represents the control group while the inner one refers to the experiment one.

Figure 3.4 *Feedback Given in High school Writing Classes*



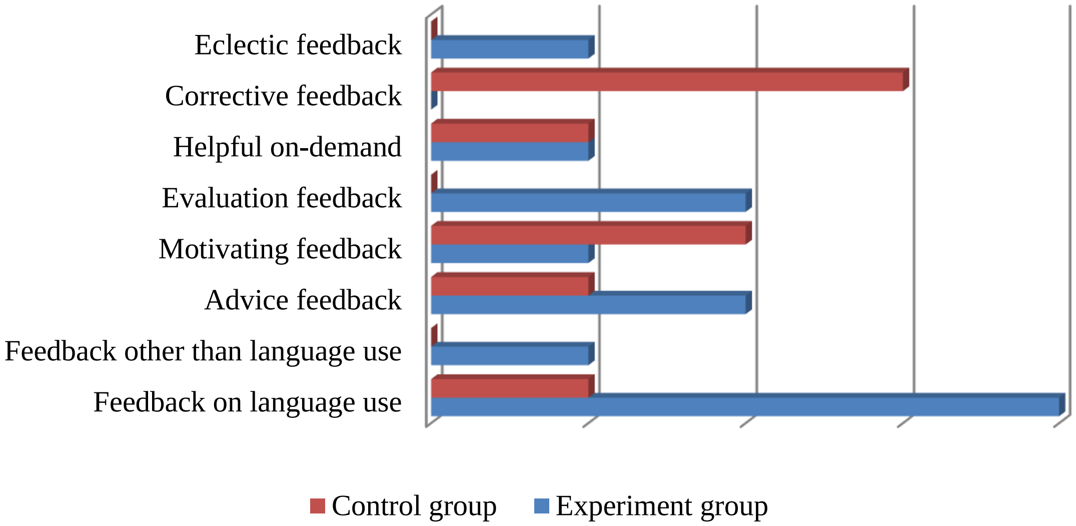
Question 1.2. Most of the participants in both groups did not like their teachers' evaluation methods followed by 'Yes' and then 'A little' answers.

Figure 3.5 *Participants' Perception of High School Teachers' Evaluation Method*



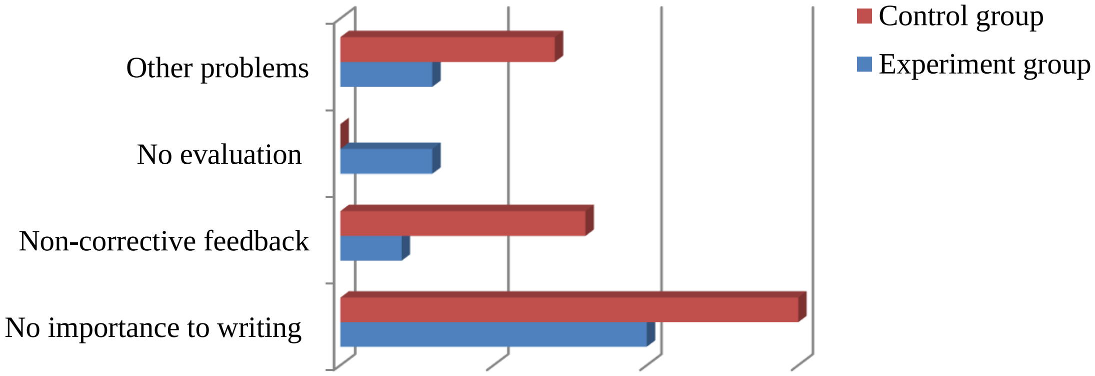
For the participants who liked their teachers' feedback, i.e. those who answered 'Yes', their answers were categorized into eight categories, as shown in Figure 3.6. The major reasons for liking their teachers' feedback in the experiment group were 'providing feedback on language use' followed by 'providing evaluation' and 'giving advice' while for the control group were 'spotting mistakes' followed by 'motivating feedback'. The three categories emerged for the experiment group, 'providing evaluation feedback', 'feedback other than language use', and 'eclectic feedback', are absent for the control group.

Figure 3.6 *Reasons for Liking the Evaluation Method ('Yes')*



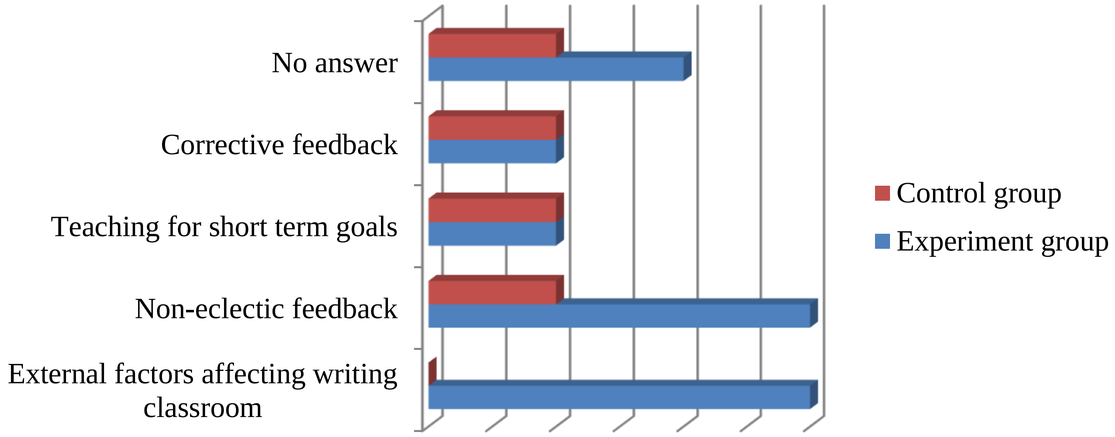
The main rationale for a 'No' answer seems to be related to the lack of priority placed on writing, as well as the absence of an evaluation for the control group.

Figure 3.7 *Reasons for Disliking the Evaluation Method ('No')*



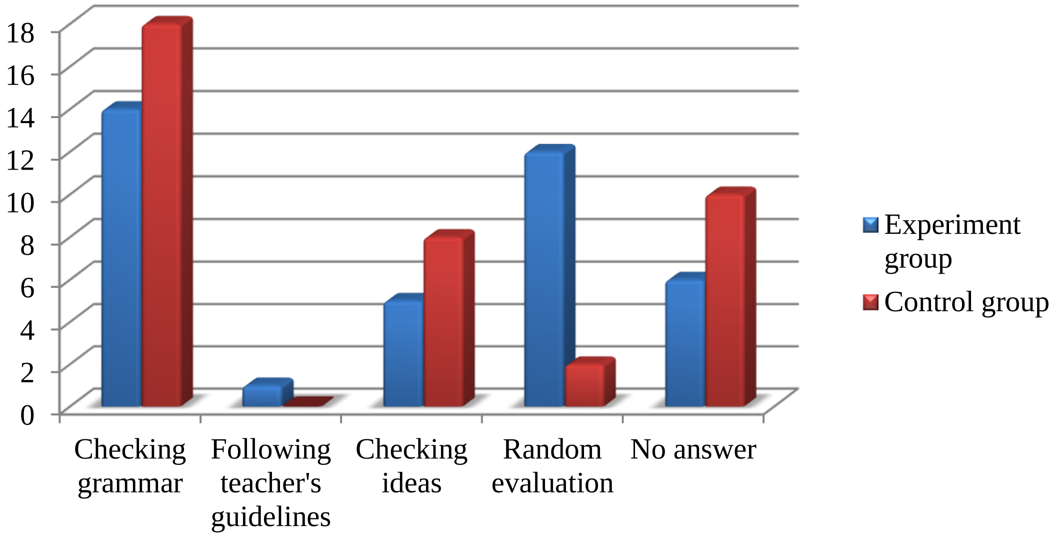
The arguments developed for 'A little' option are basically 'teaching for short term goals' and 'external factors affecting writing classroom' such as large classes and time constraints.

Figure 3.8 *Reasons for Liking the Evaluation Method ('A little')*



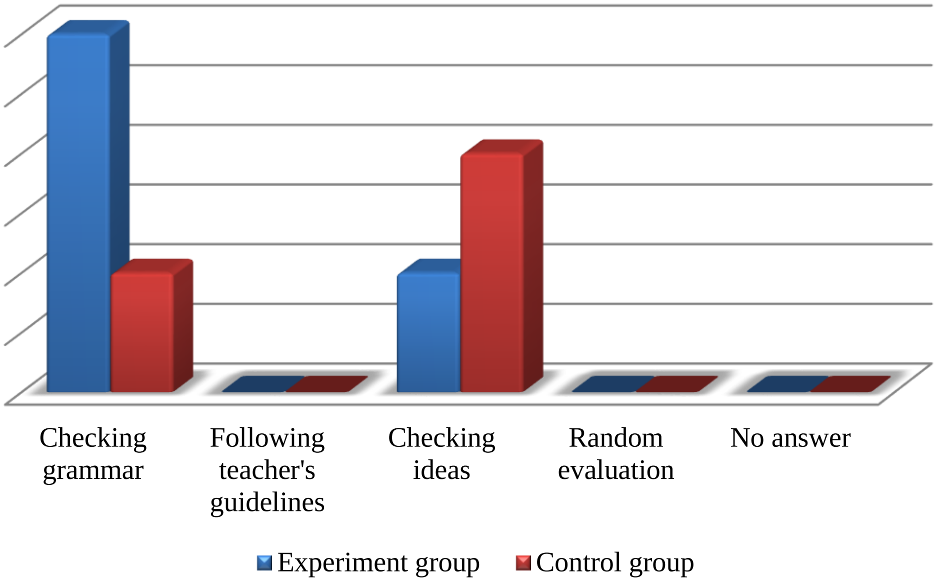
Question 2. Except for seven participants in the experiment group and 12 in the control group who completed self-assessment in high school, the other participants had no prior experience with this type of student assessment. For participants who had participated in self-assessment, the emphasis was predominantly on 'grammar use' followed by 'random evaluation'. This applies to the control group. In addition to the two preceding categories, 'checking ideas' was found a distinct category for the experiment group. When completing self-assessment, very few participants stated that they 'followed their teachers' instructions'. This category was absolutely absent in the control group.

Figure 3.9 *Self-assessment Techniques in High Schools*



Question 3. In comparison to the number of participants who experienced self-assessment in high school, doing PA was nearly absent for both groups ; except for three participants in experiment group and two in the control group. As in self-assessment, the participants were also asked to explain how they peer-assessed in case they do that. The two major categories identified for both groups were 'checking grammar' and 'checking ideas'. For the experiment group, much focus was on checking grammar and then checking ideas whereas the control group did the opposite.

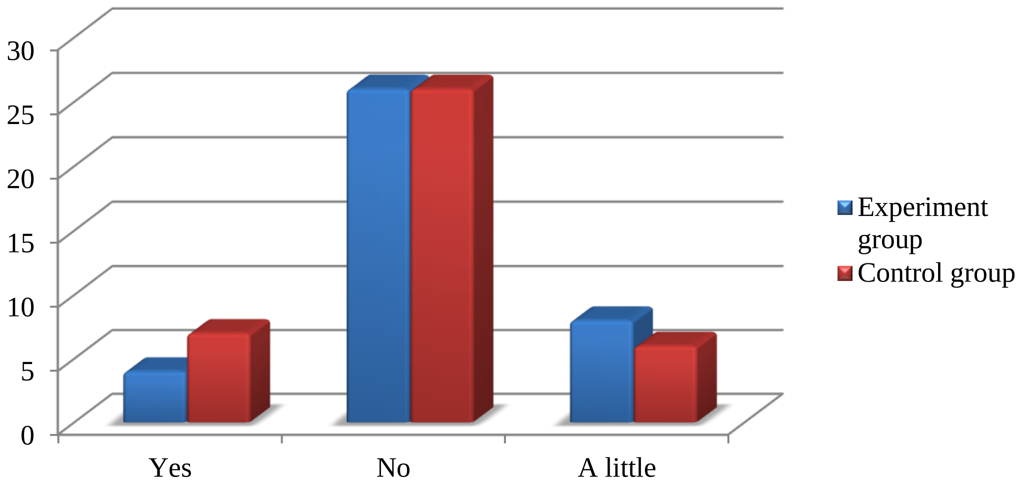
Figure 3.10 *Peer assessment Techniques in High School*



3.1.1.3 Analysis of the Third Section: Attitudes toward English Writing

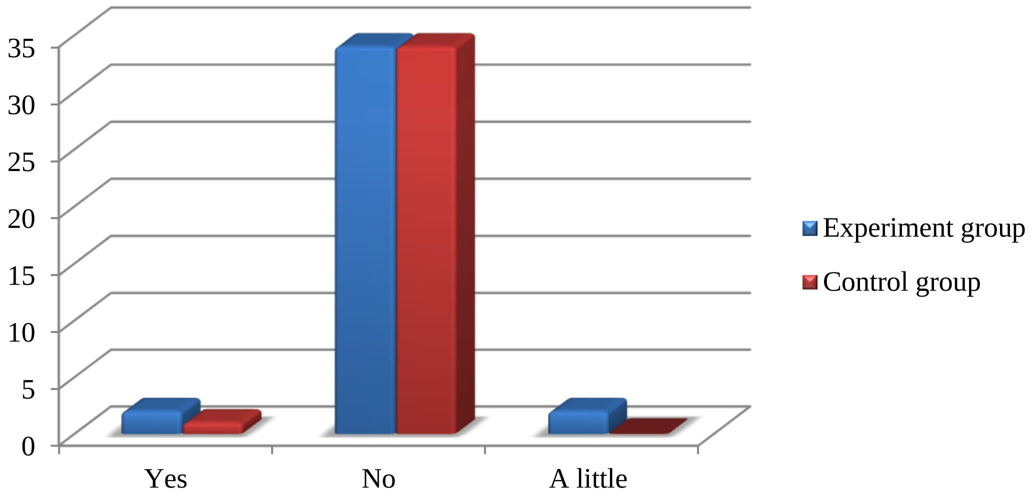
Question 1. As shown in Figure 3.11, the majority of the participants did not like writing, except few of them who answered 'Yes' or 'A little'.

Figure 3.11 *Whether Liking Writing*



Question 2. As presented in Figure 3.12, the quasi-total of participants did not like writing in their free time.

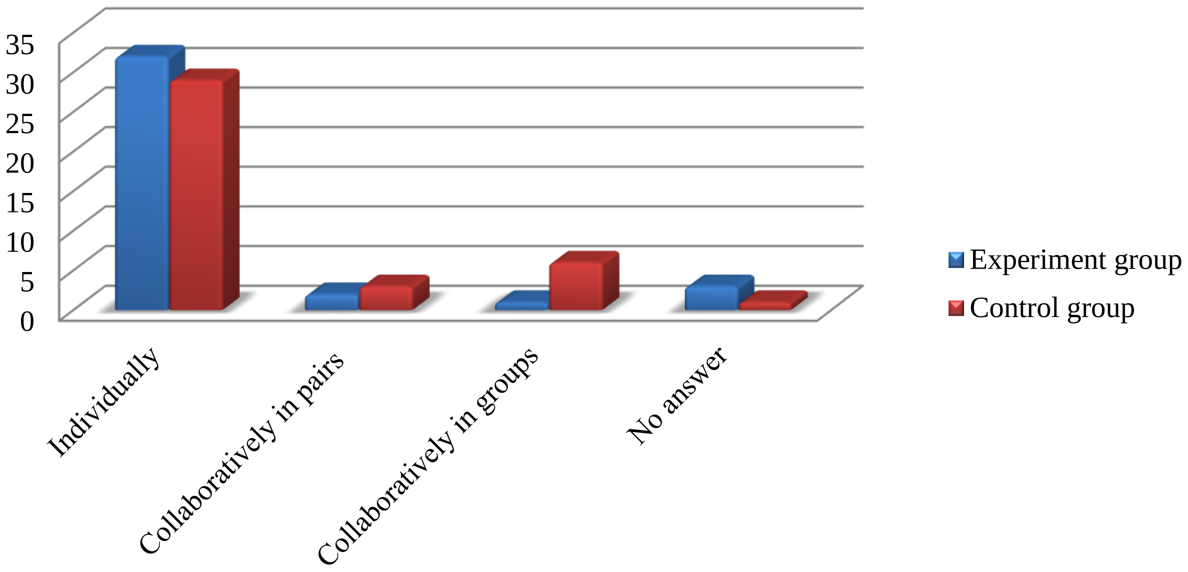
Figure 3.12 *Writing in Free Time*



Question 3. None of the participants in either group attended additional classes to improve their English writing.

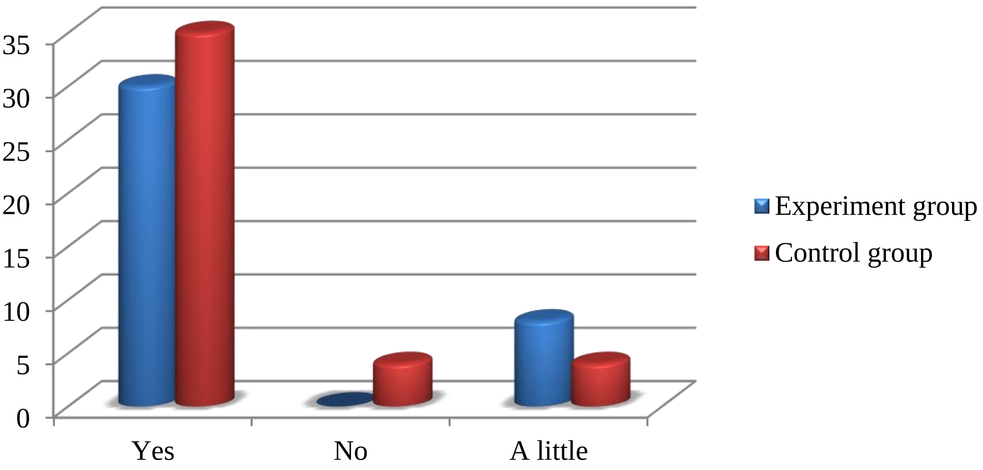
Question 4. The majority of participants preferred writing individually over writing collaboratively in groups or pairs. The number of participants who preferred writing individually was slightly higher in the experiment group than in the control one. On the contrary, writing collaboratively in groups or pairs was slightly higher in the control group than in the experiment one. All these details are presented in Figure 3.13.

Figure 3.13 *Writing Preferences*



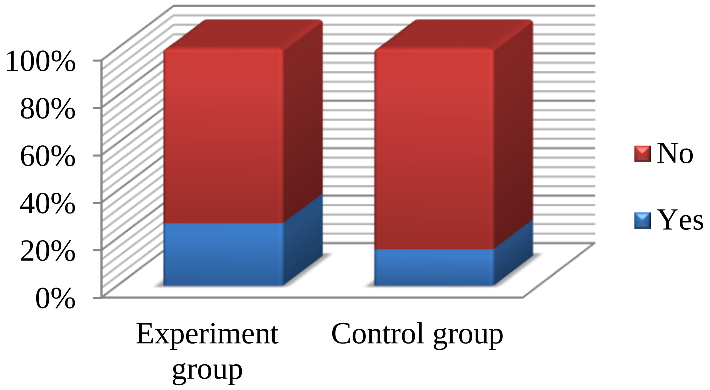
Question 5. The majority of participants in the control group, excluding four, reported feeling tense and nervous while writing. Unlike the control group, all participants in the experiment group felt worried and stressed while writing; whether they were extremely anxious or only somewhat so.

Figure 3.14 *Writing Anxiety and Stress*



Question 6. Aside from the main problems identified, such as feeling stressed and uncomfortable when writing and not being motivated to write, the majority of participants did not revise what they had written.

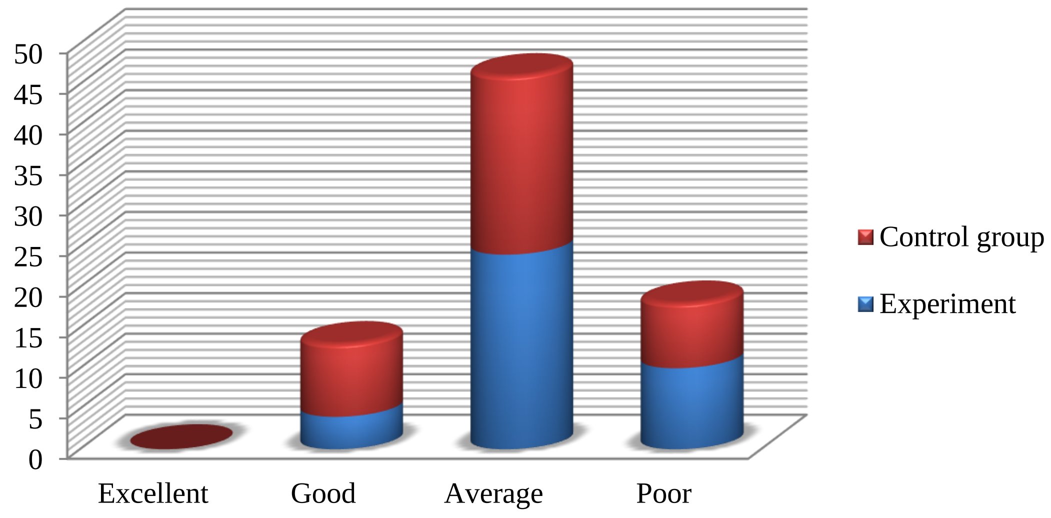
Figure 3.15 *Writing Revision*



3.1.1.4 Analysis of the Fourth Section: Beliefs about English Writing and Writing Assessment

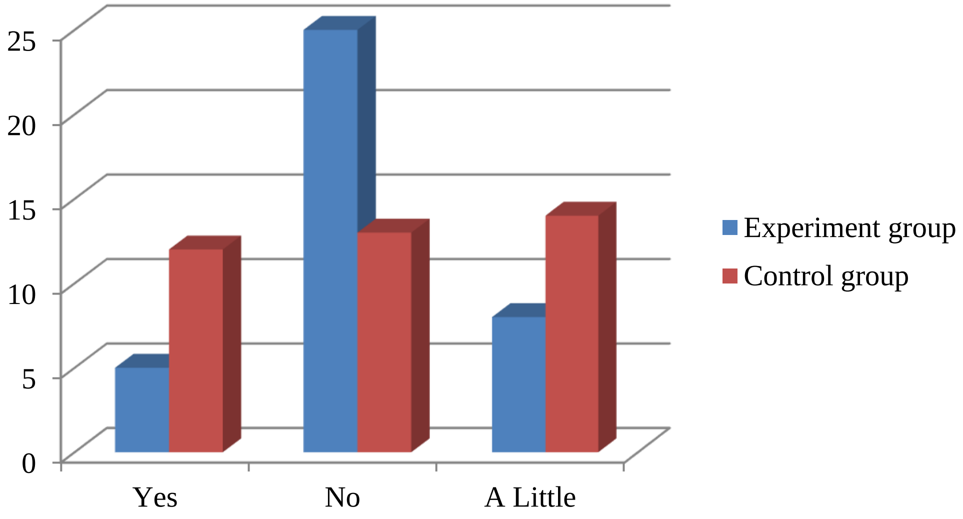
Question 1. The majority of participants in both groups thought they had an average level in writing followed by poor and good, and none thought it was excellent.

Figure 3.16 *Participants' Perception of their Writing Level*



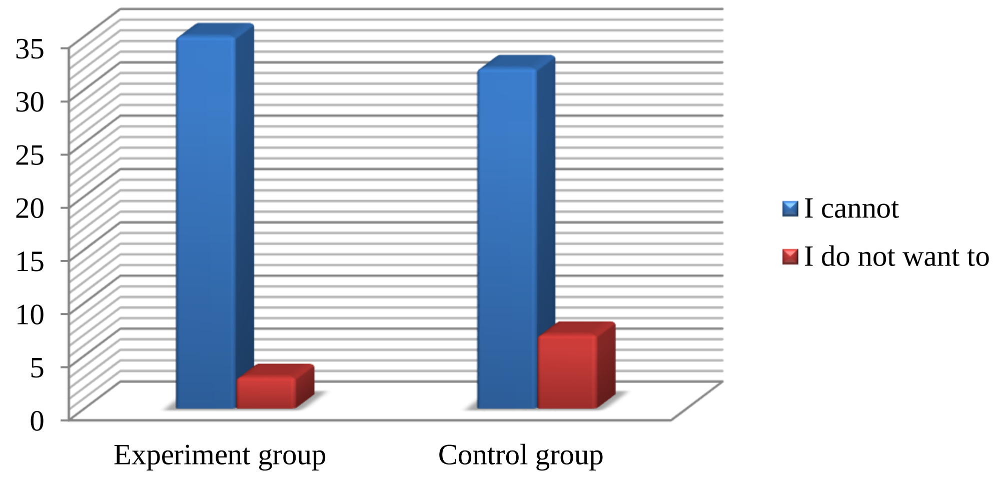
Question 2. The majority of participants stated that they lacked self-confidence in their writing.

Figure 3.17 *Participants' Perception of their Writing Self-confidence*



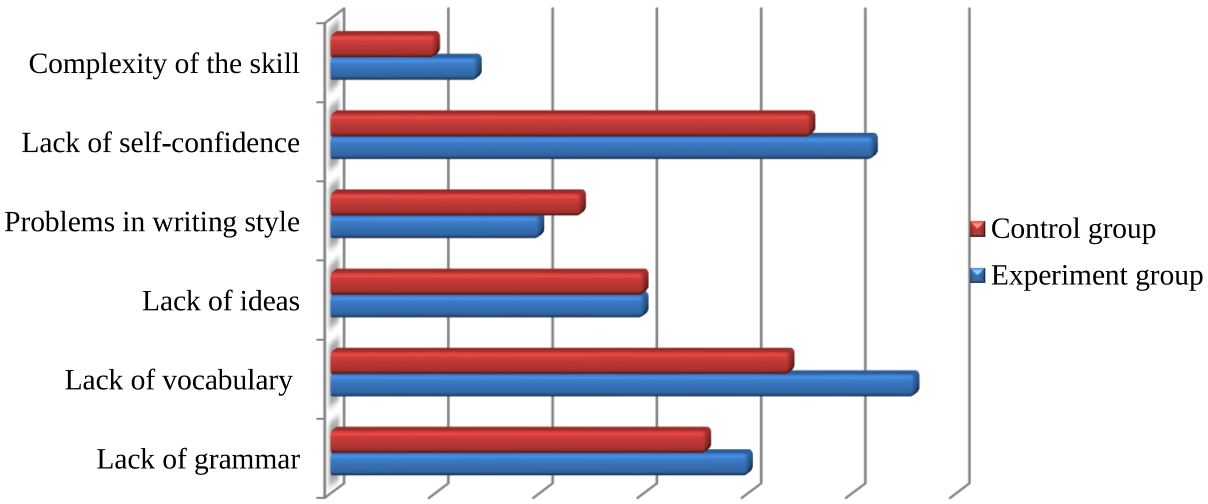
Question 3. The participants explained their poor writing as a result of their inability to write, followed by a lack of desire to do so.

Figure 3.18 *Justifying Poor Writing*



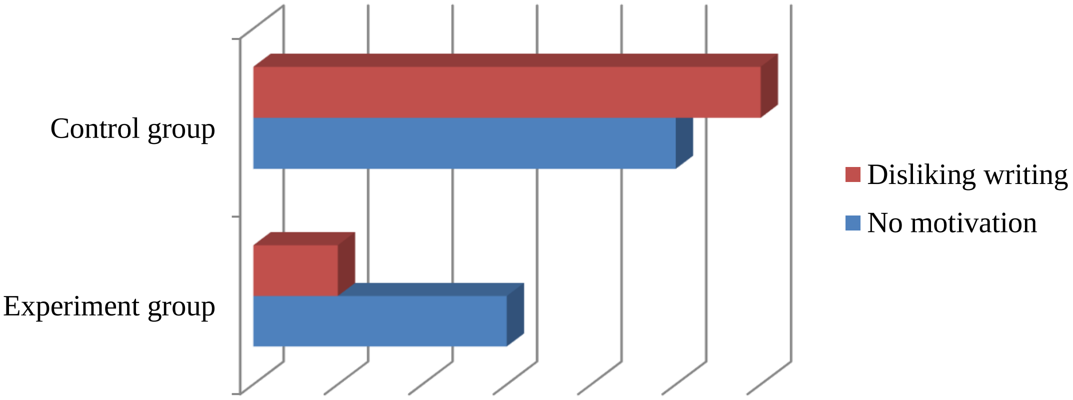
The participants explained 'cannot write well' basically with 'lack of vocabulary', 'lack of self-confidence', and 'lack of grammar'. These categories are followed by 'lack of ideas', 'writing style problems' and 'problems generated from the complexity of the skill'.

Figure 3.19 *Justifying 'Cannot' Option*



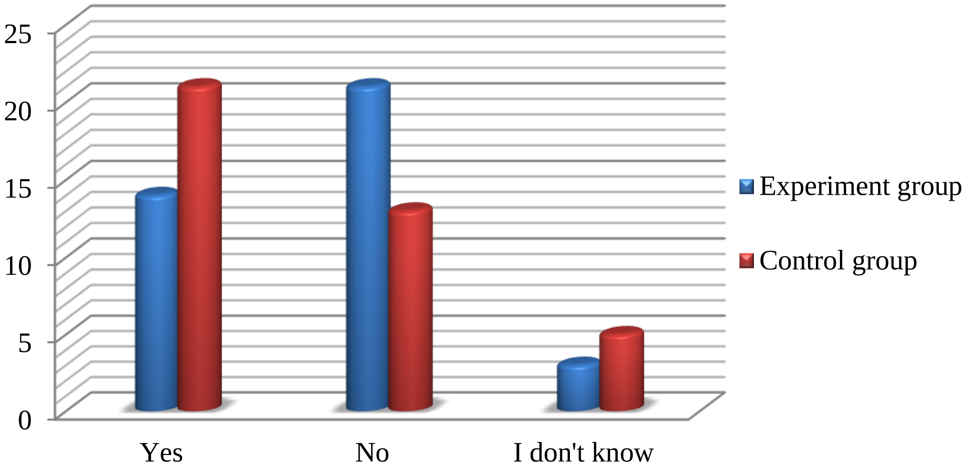
The second option was justified by two main arguments which are the absence of motivation and dislike of writing.

Figure 3.20 *Justifying 'Do not want to' Option*



Question 4. Coincidentally, the number of participants in the experiment group who thought they were not good users of grammar is the same as the ones in the control group who thought they were strong users of grammar.

Figure 3.21 *Whether Being a Good User of Grammar*



All participants who believed they were poor grammar users revealed multiple difficulties at three major levels: at the sentence level, beyond the sentence level, such as inter-sentence linkage, and additional problems, such as a lack of practice. The primary issues they faced at the sentence level were with tenses and employing conjunctions. Figure 3.22 depicts these issues in further detail.

Figure 3.22 *Overall Grammar Difficulties*

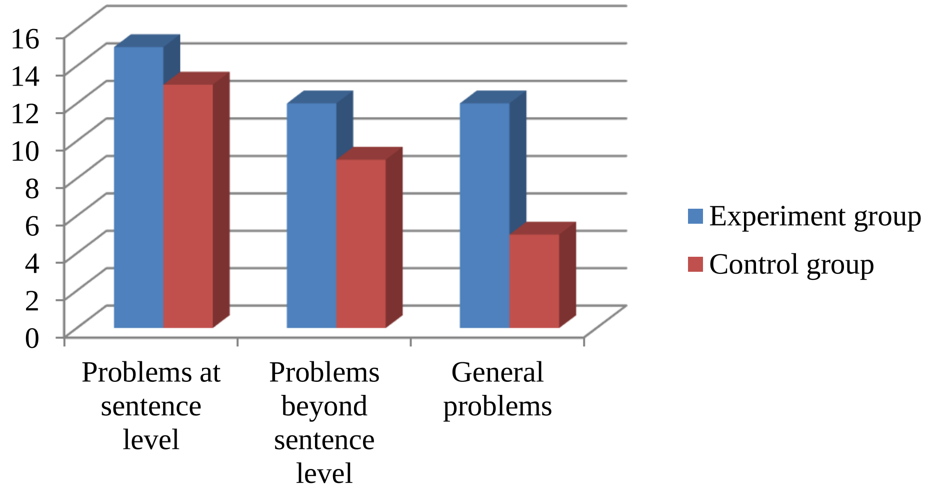
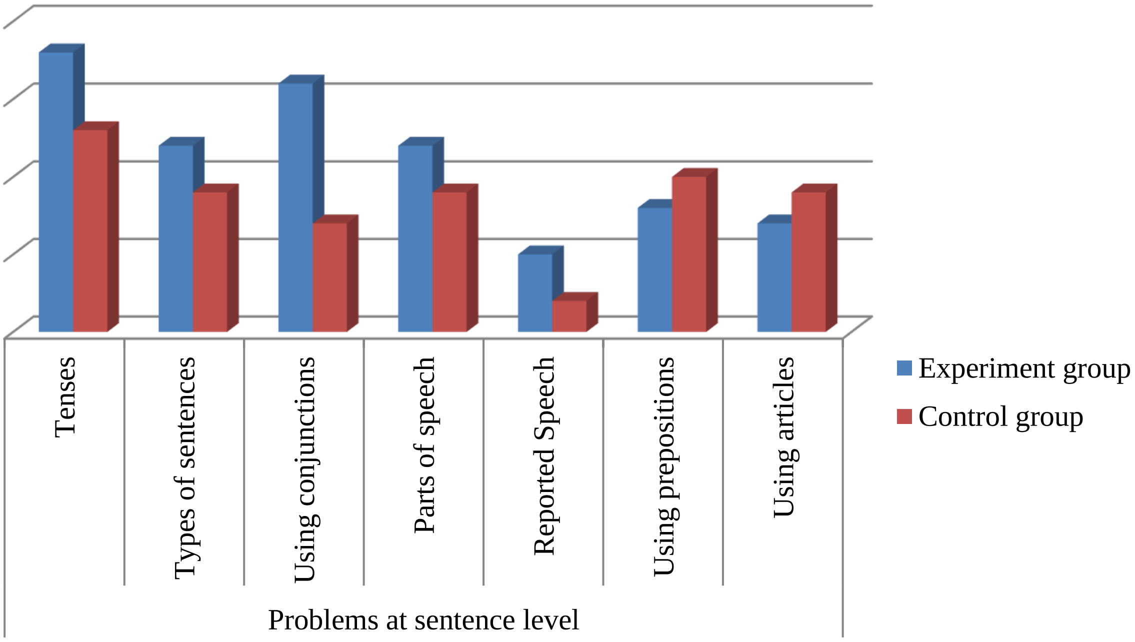
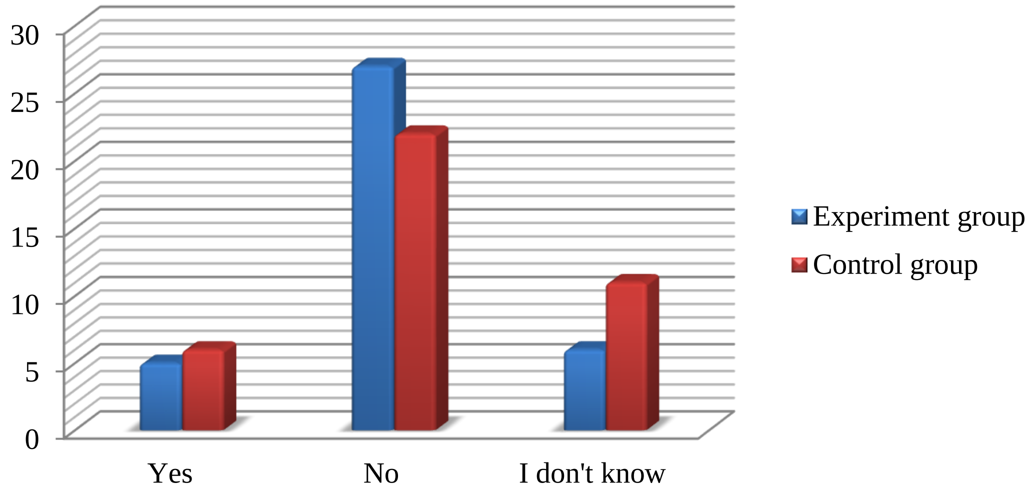


Figure 3.23 *Grammar Difficulties at Sentence Level*



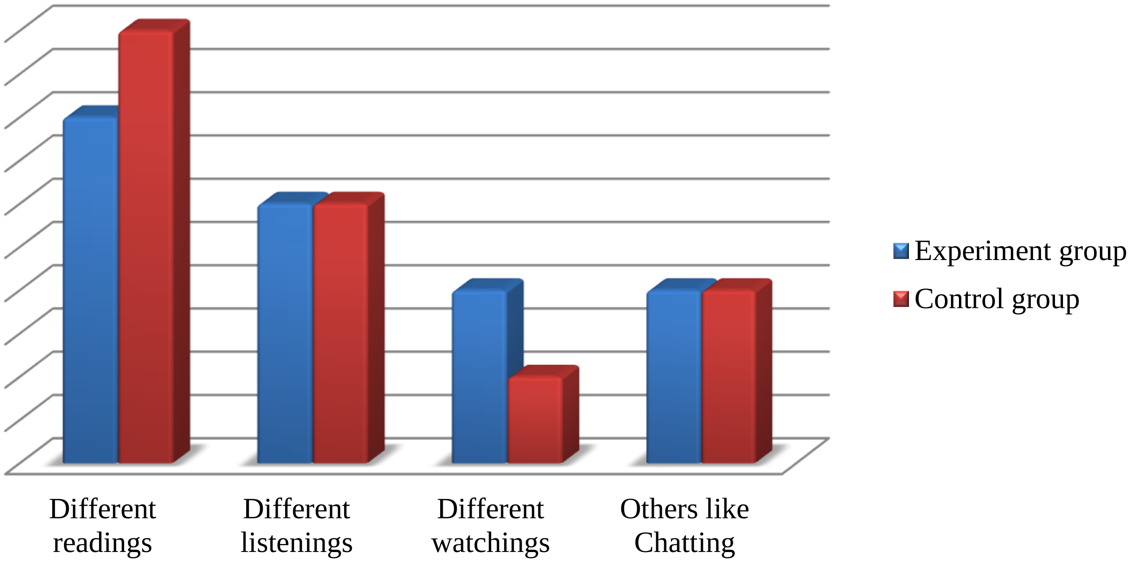
Question 5. Not similarly with grammar, both groups believed they did not have rich vocabulary.

Figure 3.24 *Whether Having Rich Vocabulary*



Question 5.1 Participants who believed they had a large vocabulary described where they gained their vocabulary. This question was addressed primarily to explore whether they liked reading over listening since the researcher asserted that reading could help them improve their writing. It is true that only reading is not enough but this skill was also included in the syllabus of teaching writing.

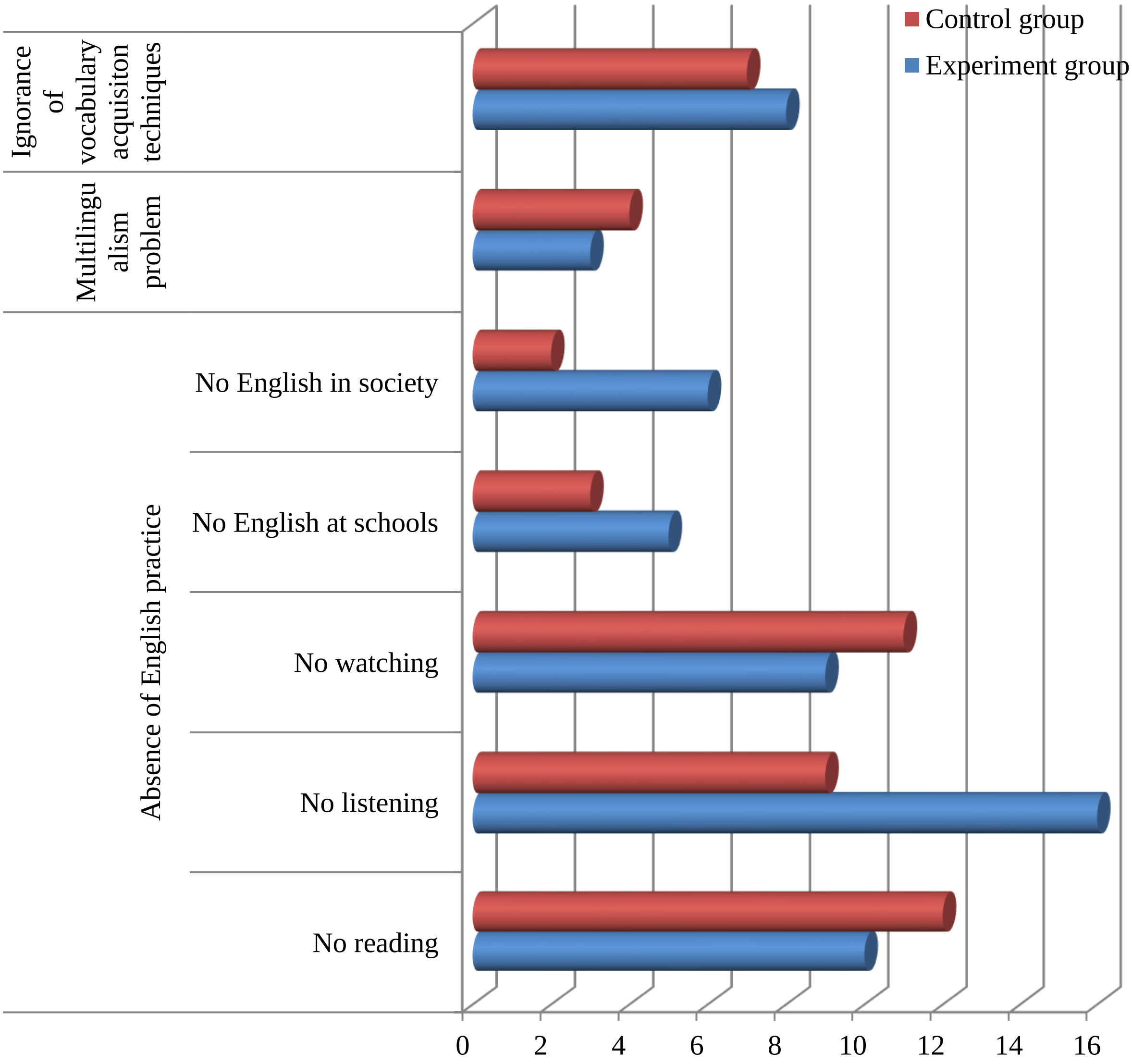
Figure 3.25 *Sources of Acquiring Vocabulary*



Question 5.2. Participants justified their low vocabulary knowledge in various ways. The causes are primarily classified into three categories: a lack of English practice,

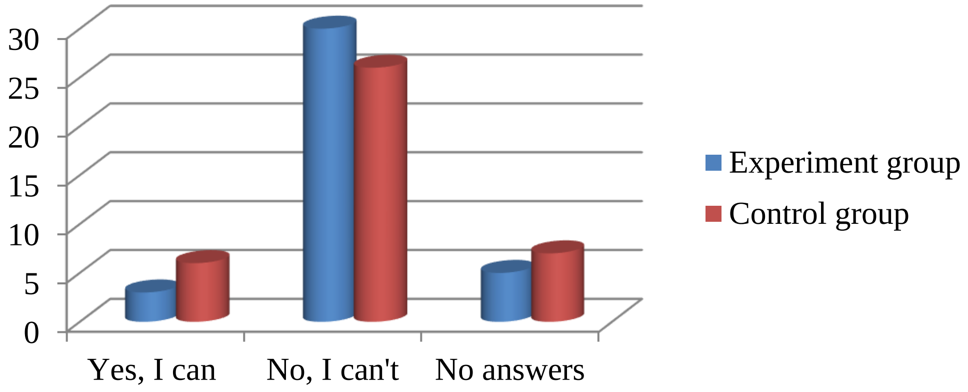
issues with multilingualism, and a lack of knowledge of vocabulary acquisition procedures. Figure 3.26 displays further details about these reasons because the first category is sub-divided into sub-categories.

Figure 3.26 *Causes of Having Poor Vocabulary*



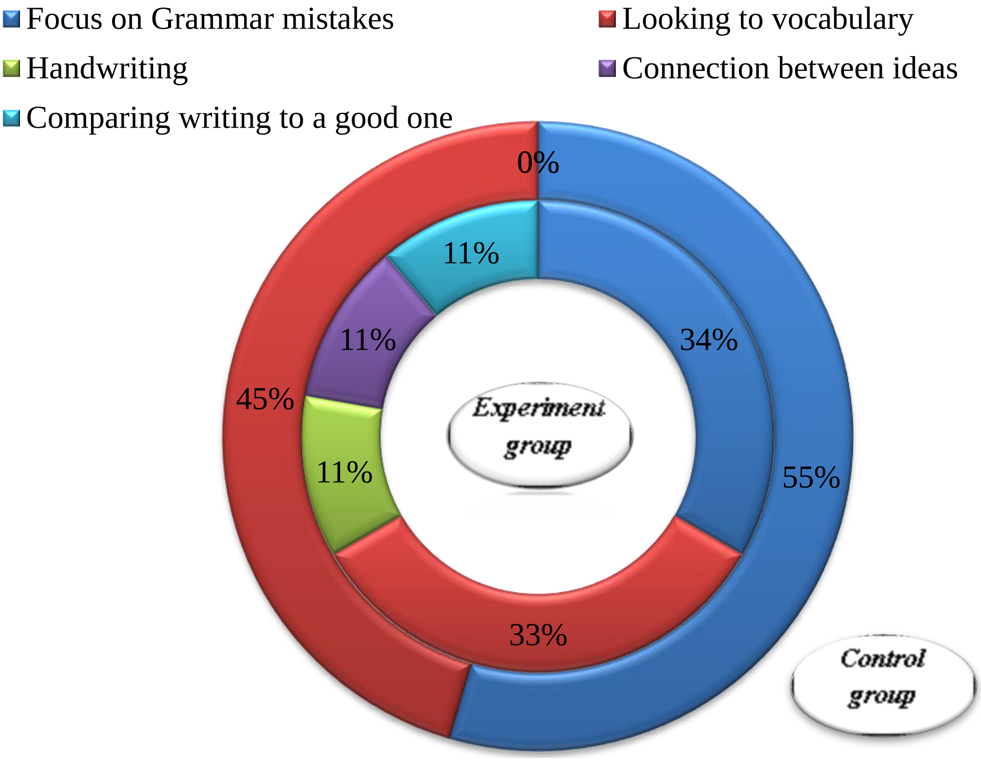
Question 6. Despite the existence of a few participants' responses who stated the opposite, the majority of them believed they could not evaluate their own writing.

Figure 3.27 *Ability to Self-assess*



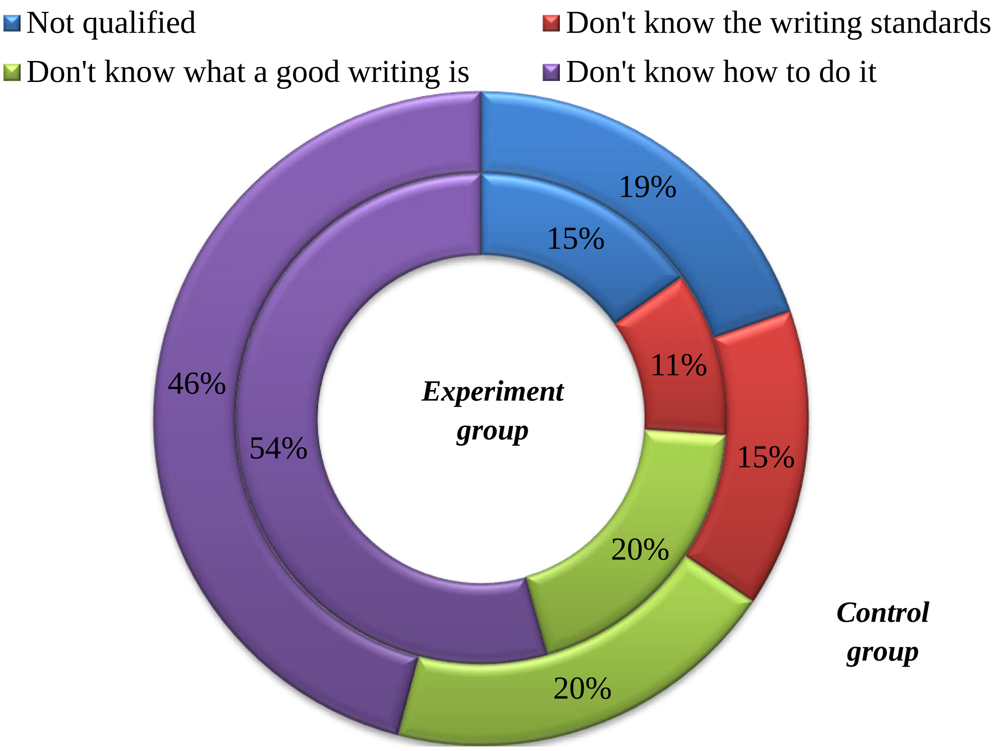
Question 6.1. The participants who responded 'Yes' for being able to self-assess were further asked to describe how. Their answers were limited to focusing on grammar and vocabulary for the participants in the control group while in the experiment group the participants suggested other reasons as displayed in Figure 3.28.

Figure 3.28 *Explaining Ability to Self-assess*



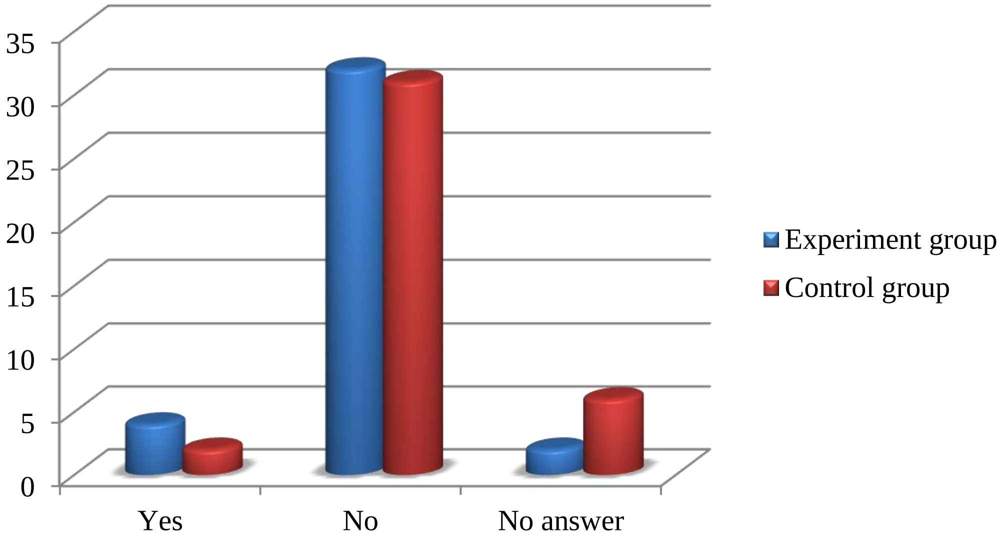
Question 6.2. The participants explained their inability to self-assess their writing with four major reasons, as it is shown in Figure 3.29.

Figure 3.29 *Explaining Inability to Self-assess*



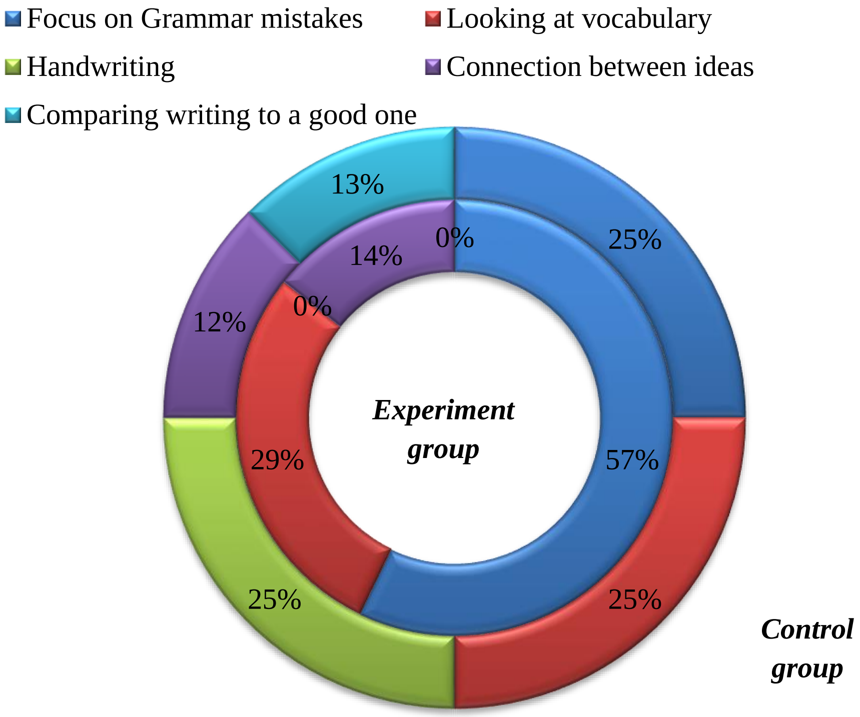
Question 7. As with peer assessment, the majority of participants more than in self-assessment, said they were unable to peer appraise their classmates' works.

Figure 3.30 *Ability to Peer assess*



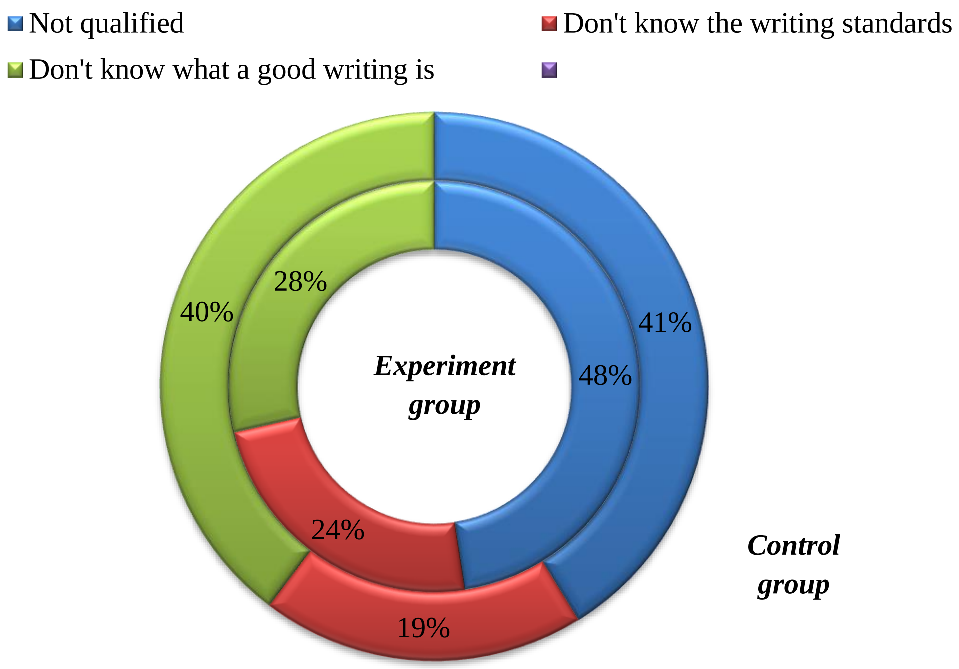
Question 7.1. The participants who thought they could peer assess were then asked to explain how they could do it. As Figure 3.36 displays, the participants in the experiment group thought that peer assessment happens when focusing on grammar mistakes (this point was highly emphasized), looking at good vocabulary, and checking connection between ideas. The participants in the control group added comparing writing to a good one and looking at connection between ideas.

Figure 3.31 *Explaining Ability to Peer Assess*



Question 7.2. Participants who believed they could not peer assess their classmates' work indicated that they did not have the necessary skills or knowledge of the writing standards.

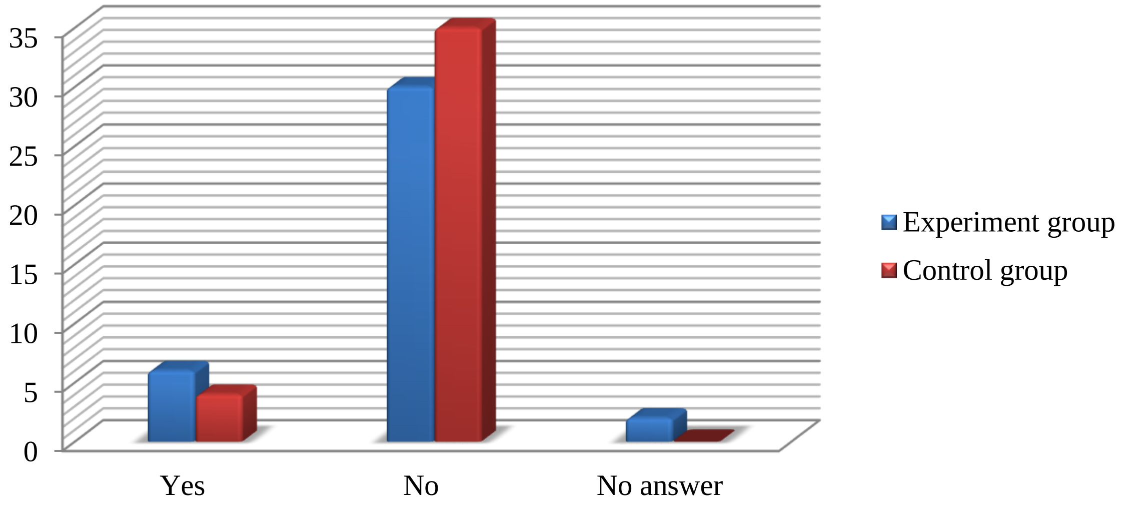
Figure 3.32 *Explaining Inability to Peer Assess*



3.1.1.5 Analysis of the Fifth Section: Needs and Awareness toward English Writing

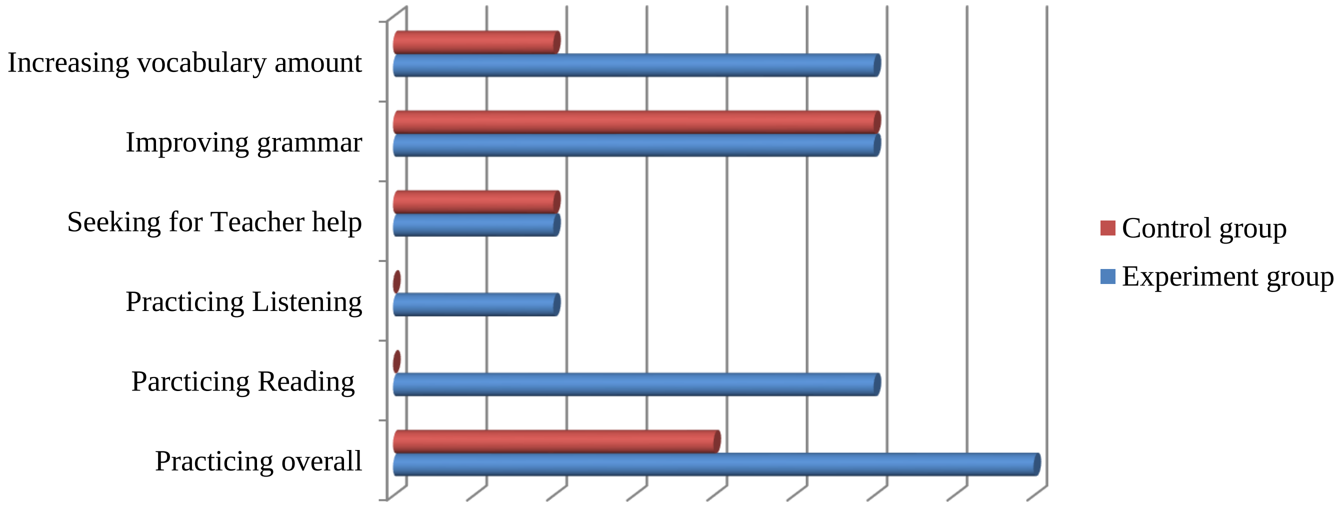
Question 1. Nearly, all participants stated that they had no idea how to enhance their writing.

Figure 3.33 *Participants' Ability to Improve their Writing*



When asked whether they can improve their writing, the participants in the experiment group suggested primarily ‘practicing writing’, ‘practicing reading’, ‘improving grammar’, ‘increasing vocabulary’, ‘teacher help’ and ‘practicing listening’ while in the control group the participants suggested just ‘improving grammar’, ‘increasing vocabulary’, ‘practicing’ and ‘teacher help’.

Figure 3.34 *Suggested Ways by Participants to Improve Writing*



Question 2. Almost all participants wanted to talk individually with their teacher for many reasons exposed in Figure 3.35. The reasons for no need of an individual interaction student-teacher are exposed in Figure 3.36.

Figure 3.35 *Face to Face Student-Teacher Interaction*

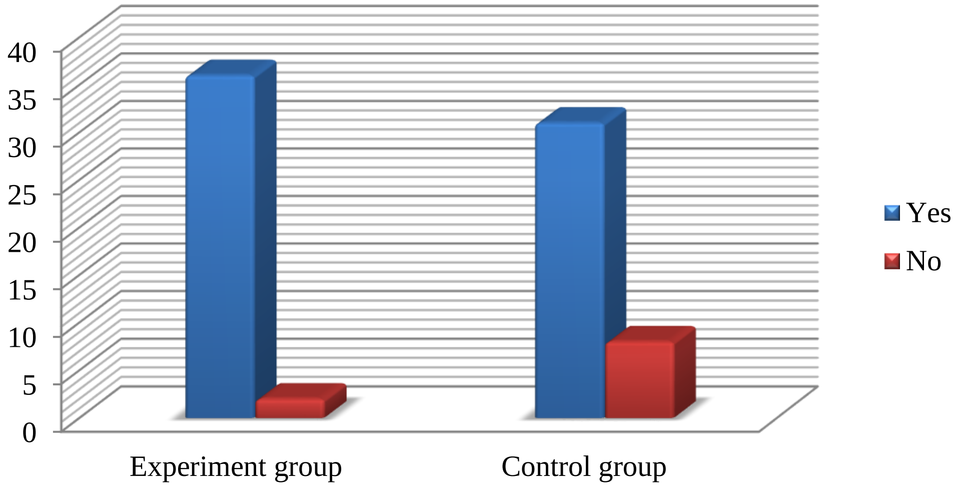


Figure 3.36 *Reasons for the Need of Student-teacher Interaction*

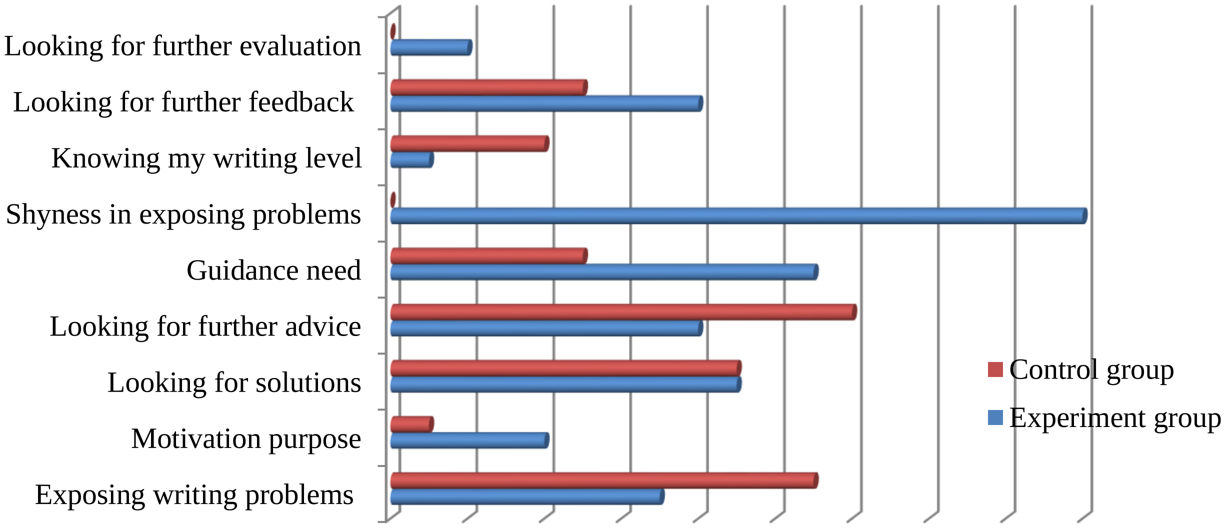
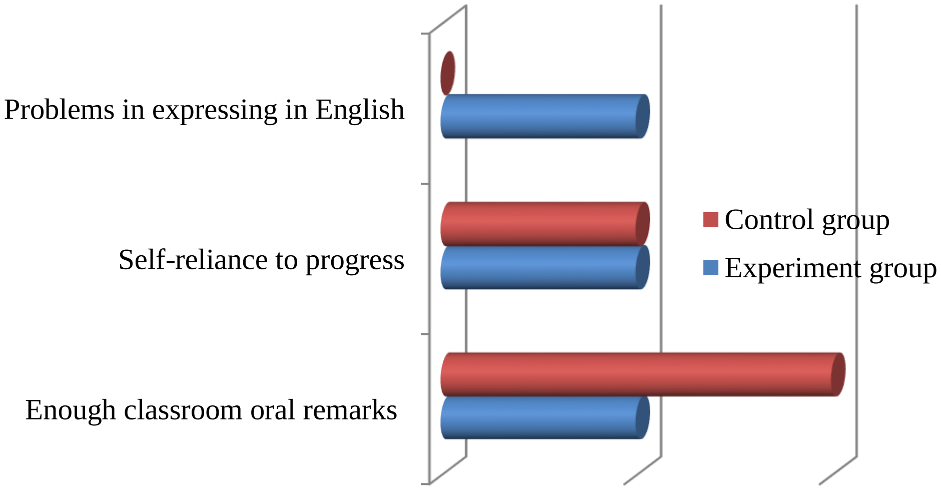
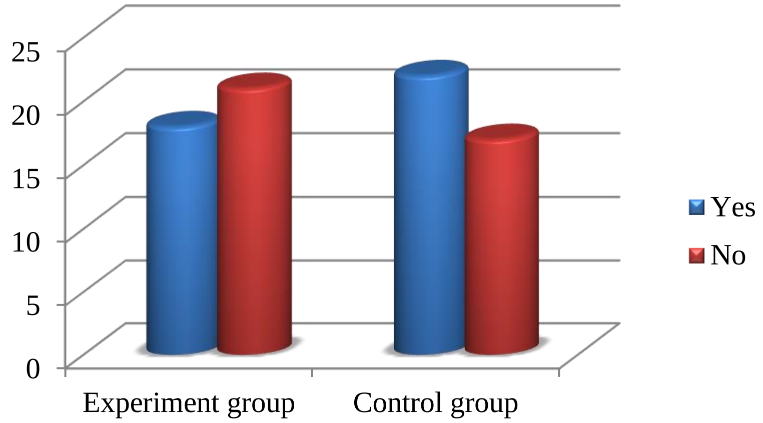


Figure 3.37 *Reasons for No Need of Student-teacher Interaction*



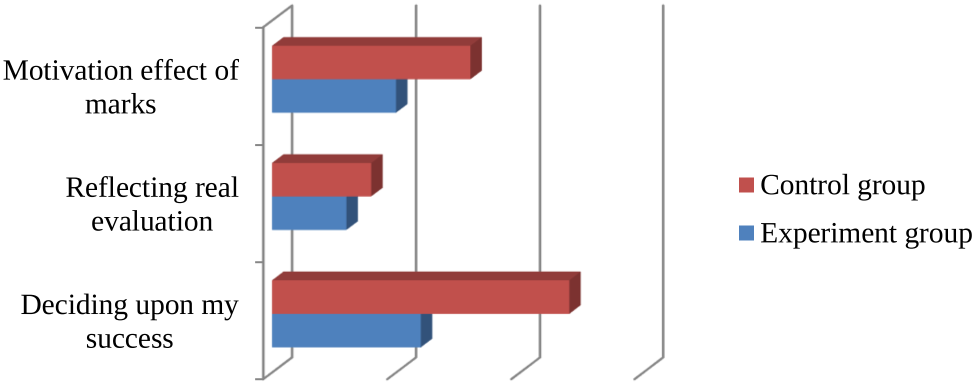
Question 3. Participants' need for feedback was slightly higher than their need for marks. It means both, feedback and marks, are equally important for them.

Figure 3.38 *Whether Preferring Marks over Feedback*



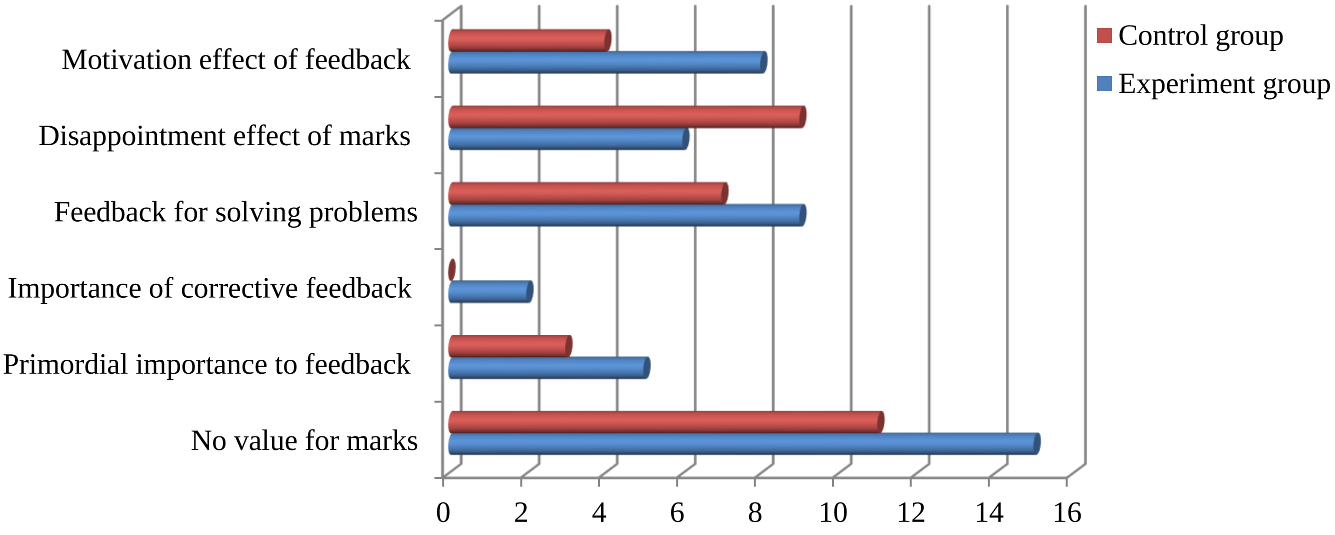
Those who said that marks are more important than feedback suggested that marks are key to success and motivation and represent the real evaluation.

Figure 3.39 *Reasons for Praising Marks over Feedback*



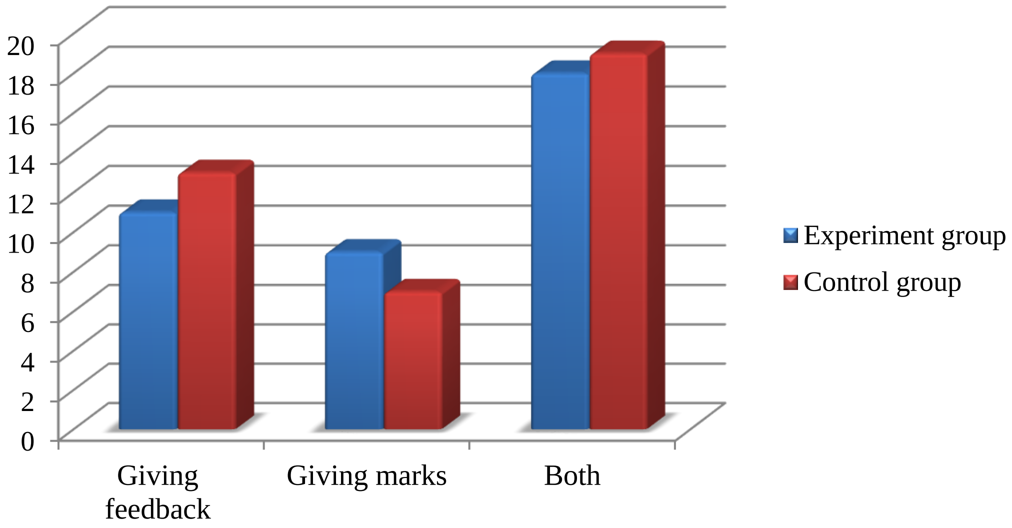
Those who believed that feedback was preferable to grades defended their position differently, as displayed in Figure 3.40.

Figure 3.40 *Reasons for Praising Feedback over Marks*



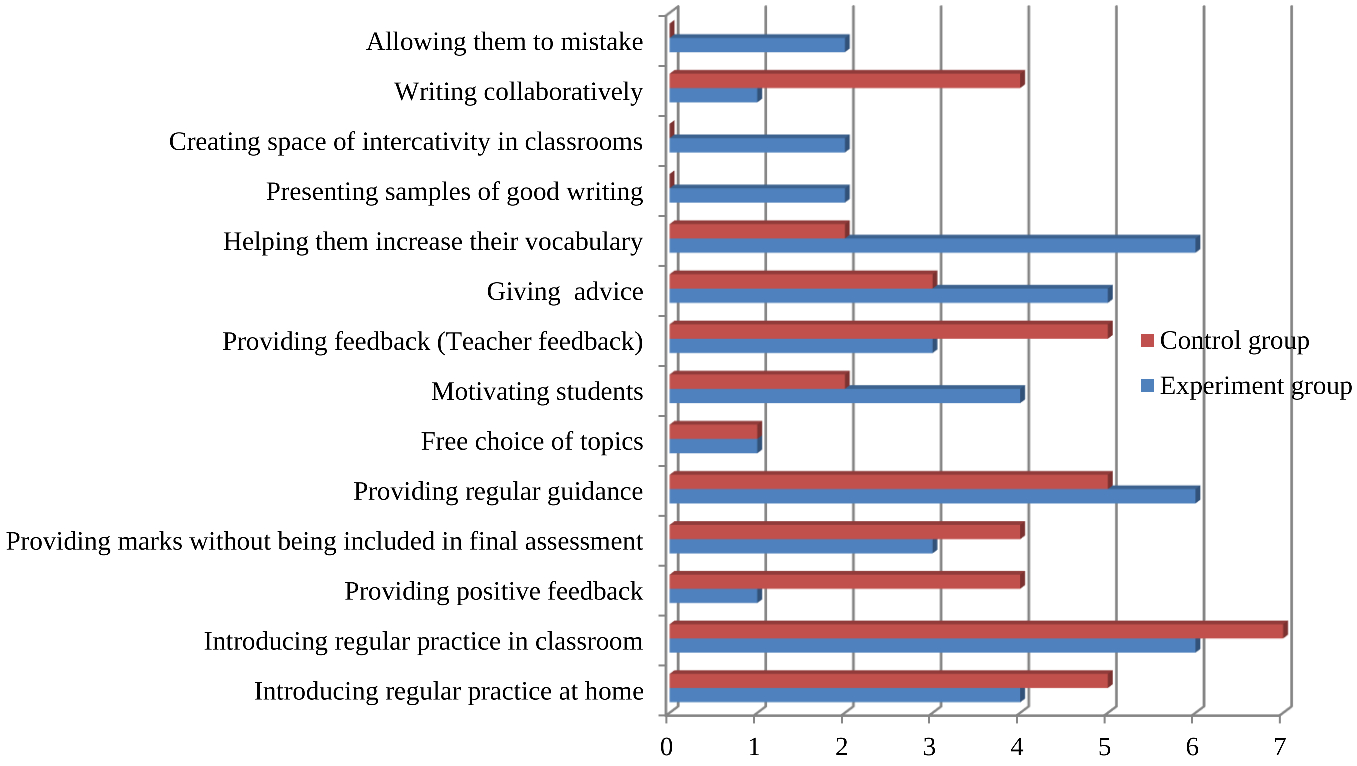
Question 4. The participants suggested combining both marks and feedback, followed by giving feedback, and finally giving marks.

Figure 3.41 *Participants' Suggestion for the Need for Marks, Feedback, or Both*



Question 5. This question was asked purposively to put participants in 'the shoes of the teacher' so that they could be aware of teaching writing and how it might be taught effectively to students. This might help in providing helpful ideas and raising their awareness to be responsible of their own writing progress.

Figure 3.42 *Participants in Teacher's Position*



3.1.2 Analysis of the Written Pre-test

The purpose of analyzing the pre-test is to identify different errors in participants' writings based on which the rubrics should be designed. Toward this aim, 168 samples of four groups of the population study were gathered. Their analysis has been conducted in accordance with James's (2013) taxonomy, with the inclusion of a category of difficulties relating to paragraph form. This taxonomy deals with the errors in language learning, including the four skills, classifying them into: substance errors, textual errors, and discourse errors. In writing, the errors are called respectively: "misspelling", "miswriting", and "miscomposing" (p. 130). Table 3.1 below displays the different errors and problems spotted in the writings of the participants with their corresponding percentages for the four groups.

The errors and problems found in the participants' writings were highlighted without judging the written piece as poor, average, good, or excellent. The researcher indicated just the presence or not of the corresponding error or problem because the aim was to spot and not to evaluate in this stage of research. It is obvious that an excellent or a perfect written piece is free of errors and problems. Consequently, it is worth bearing in mind that ‘present’ did mean the corresponding error or problem was present anywhere in the written piece, excluding any interpretation that it was everywhere. This indicates that the presence of the related errors and problems in the written work does not rule out the possibility of the writing to be good or even exceptional, as this can only be determined through explicit assessment criteria such as using assessment tools, namely rubrics, prompts, or checklists. Percentages describe the frequency of the problem or inaccuracy in the entire sample, not the frequency of its occurrence in each individual piece. The slash in the table indicates that the component was not specified with explicit rules. For example, in capitalization, all of the requirements for appropriate capitalization must be followed without exception.

3.1.2.1 Substance Errors. According to James (2013), substance errors are errors that involve the phonological or the graphological substance systems. This kind of errors is primarily linked to mechanics, which are punctuation errors, typographic errors, dyslexic errors, and confusibles. For the current study, the errors spotted are spelling, punctuation, and capitalization errors. Basically, for spelling errors Kusuran (2016) sees that they can be categorized into two main categories: typographic errors and cognitive errors. Typographic errors include substituting, omitting, adding, or transferring letters, but cognitive errors include letters which have phonetic similarities like ‘academic’ and ‘akademic’. Whether being typographic or cognitive, as these errors often change sentence meaning, the researcher of this study considered all the spelling errors cognitive.

3.1.2.2 Textual Errors. This category is mostly associated with misapplication of the lexico-grammar to achieve a texture (James, 2013). Textual errors that arise in this research are mostly lexical, grammar, and syntax errors. Lexical errors included lack of vocabulary, contextual meaning, and using foreign words. Analyzing grammatical errors can help teachers in teaching through designing appropriate materials and students recognize their mistakes (Hasyim, 2002). The grammatical errors included subject-verb agreement, tenses, use of articles, use of preposition, and use of negation. Finally, syntax errors are sentence errors and inter-sentence errors (cohesion errors).

For cohesion, based on the work of Halliday and Hasan (1976), five types of cohesive links are identified: reference, substitution, ellipsis, conjunction and lexical cohesion. This classification is highly relied upon in treating problems of cohesion in the EFL context, and it is also important to point to the fact that cohesive markers are not obligatory and, in some cases, even undesirable (James, 2013).

3.1.2.3 Discourse Errors. According to James (2013), this category includes coherence errors, pragmatic errors, and receptive errors. In this study, the researcher spotted only coherence errors to be the focus of this research. Coherence is viewed differently, but it is basically related to the meaning of the written text. For Bhatia (1974), coherence includes relevance, clarity, development, and even originality. However, for Das (1978) coherence is ‘value-as-message’, while ‘value-as-text’ is cohesion, which is defined in terms of communicative function involving both reader’s interpretation and writer’s intention. For further understanding, Widdowson (1995) suggests that text vs. discourse is distinguished as product vs. process and meaning vs. interpretation arguing that, “it is your discourse you read into my text” (p. 165). The coherence problems found in participants’ writings of this study were of topical and relational coherence types. Coherence errors are also called content organization errors (Ruegg & Sugiyama, 2013).

3.1.2.4 Paragraph Structure Problems. In this category, the researcher spotted the presence or not of the problems in the major components of a paragraph: topic sentence, supporting sentences, and concluding sentence. She commonly took into account that a paragraph has to include all the three components placing basically the topic sentence at the beginning of the paragraph, to be the first one, and the concluding sentence at the end, or to be the last one. The researcher excluded any exceptional cases, such as not having a topic sentence or placing it elsewhere in the paragraph. The same happened for concluding sentence; one concluding sentence places at the end.

3.1.2.5 Paragraph Form and Appearance Problems. This category was created by the researcher as she noticed significant problems in forms of the writings delivered. To ensure an appropriate form, the following items should be taken into consideration: paragraph length, paragraph entitlement, block format, indentation, and handwriting.

The importance of the last item is explained differently by some scholar. For example, McFarland (2015) sees that writing is empowering the brain, memory, motor skills, and reading. In relation to reading, Trafford and Nelson (2003) have also demonstrated the importance of handwriting to reading fluency because the illegible one is difficult to decipher.

Based on the explanation provided to categorize participants' problems in paragraph writing, Table 3.1 displays all the categories and sub-categories discussed above.

Table 3.1 *Errors and Problems Spotted in Participants' Writings*

Errors and problems			Presence			
Type	**Category**	**Sub-category**	**Exp. group**	**Cont. group**	**Group 3**	**Group 4**
Substance	Spelling	Typographic & cognitive	91%	72%	82%	62%
	Punctuation	/	100%	100%	100%	100%
	Capitalization	/	78%	72%	80%	67%
Textual	Lexical	Lack of vocabulary	82%	52%	78%	47%
		Contextual meaning	48%	43%	56%	73%
		Foreign words	05%	10%	12%	02%
	Grammatical	/	100%	100%	100%	100%
	Syntax	Sentence errors	62%	70%	92%	61%
		Inter-sentence errors (cohesion)	72%	80%	79%	62%
Discourse	Coherence	Topical coherence	74%	80%	80%	62%
		Relational coherence	62%	58%	49%	49%
Paragraph Structure Problems	Topic sentence		100%	100%	100%	100%
	Supporting Sentences		71%	85%	58%	73%
	Concluding Sentence		72%	78%	67%	80%
Form and Appearance Problems	Paragraph Length		82%	48%	62%	73%
	Paragraph entitlement		87%	90%	96%	80%
	Block format		88%	78%	52%	64%
	Indentation		78%	61%	91%	70%
	Handwriting		04%	02%	01%	03%

As shown in Table 3.1, all participants in four groups (experiment, control, and two other groups of the same population) have significant problems related to paragraph writing, except for two minor problems which are using foreign vocabulary and handwriting. The former is related to the content and the other to the form and appearance of the paragraph. The dominant errors, which were present in all the written pieces, are using punctuation marks and grammar, and problems with topic sentence. The other errors and problems are present with various percentages, but all remain significant.

3.2 Data Analysis and Presentation: Second Stage

In this stage, the data analyzed were obtained from the use of rubrics with their follow-up sections, which included self- and APA associated with peer feedback, and the researcher journal report.

3.2.1 Analysis of the Researcher Journal Report

The major aspects under concern are participant writing time, time devoted for performing self-assessment and APA, writing rubric use, and overall classroom assessment.

3.2.1.1 Participant Writing Time. The researcher chose to record participants' writing time in the experiment group over the course of twelve sessions in order to track it. The time is divided into intervals. Numbers (1, 2, and 3) show the number of times the descriptive and narrative genres were written; each was rehearsed three times.

Figure 3.43 *Participant Writing Time*

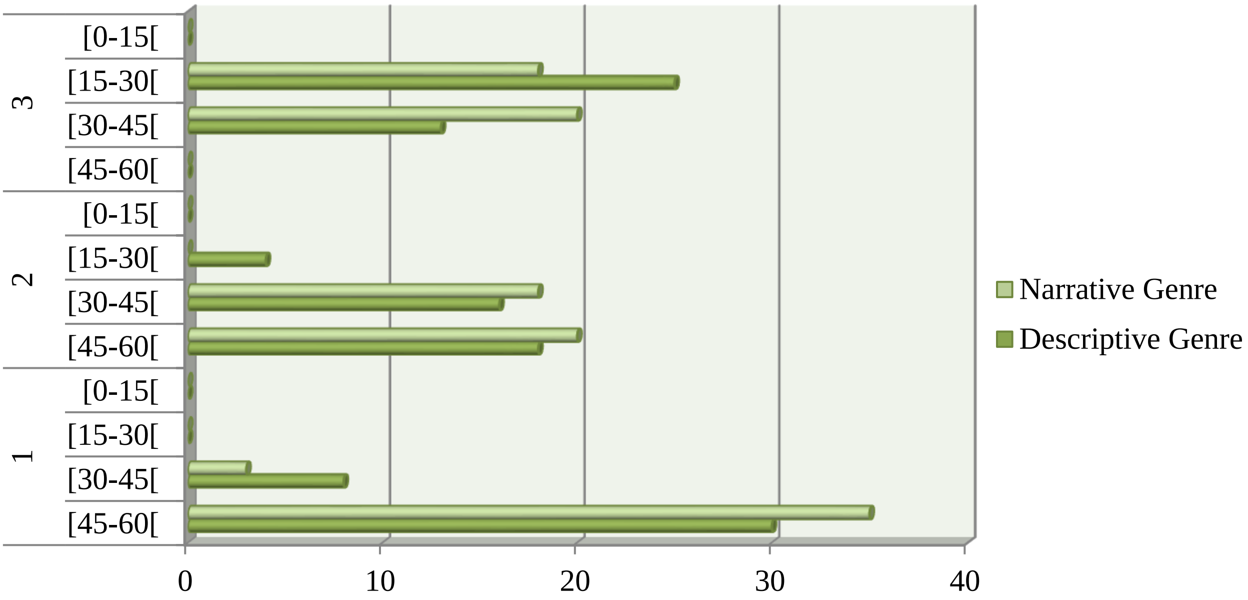
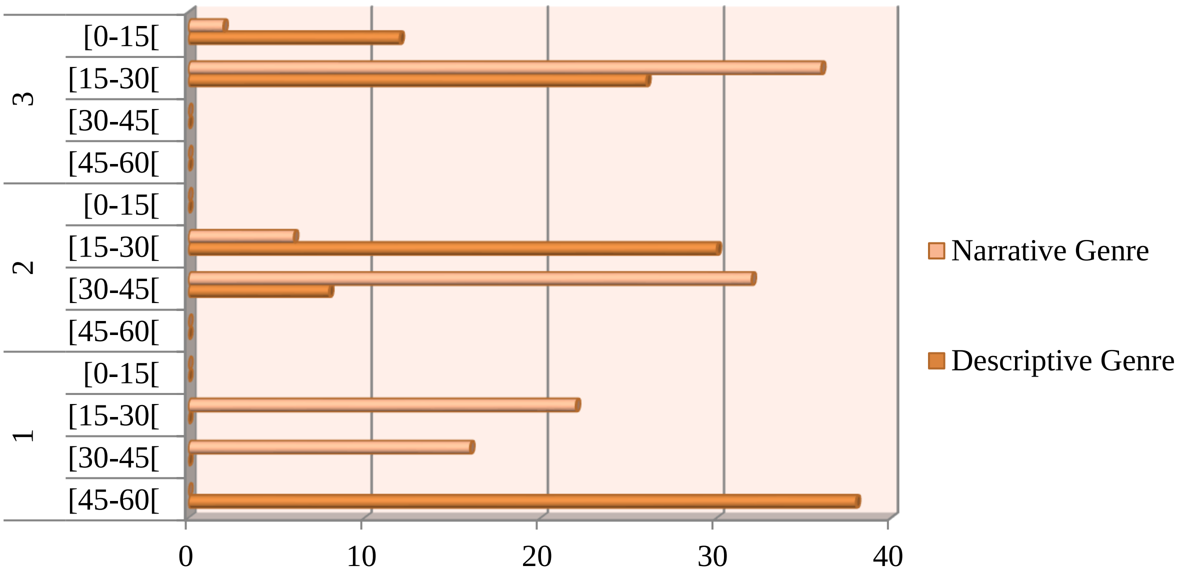


Figure 3.43 demonstrates that at the beginning the majority of participants devoted between 45 to 60 min to write a paragraph, with a higher number when writing the narrative genre. By the end of the intervention, the majority of participants had succeeded in writing the descriptive genre within 15 and 30 min as compared to the number of participants who completed that for the narrative genre in the same interval. In other words, the participants needed much more time to write the narrative genre.

3.2.1.2 Time Devoted to Self-assessment Performance. The researcher tracked the time spent completing self-assessment of both descriptive and narrative genres for comparison sake. Figure 3.44 presents the time spent over all the practices.

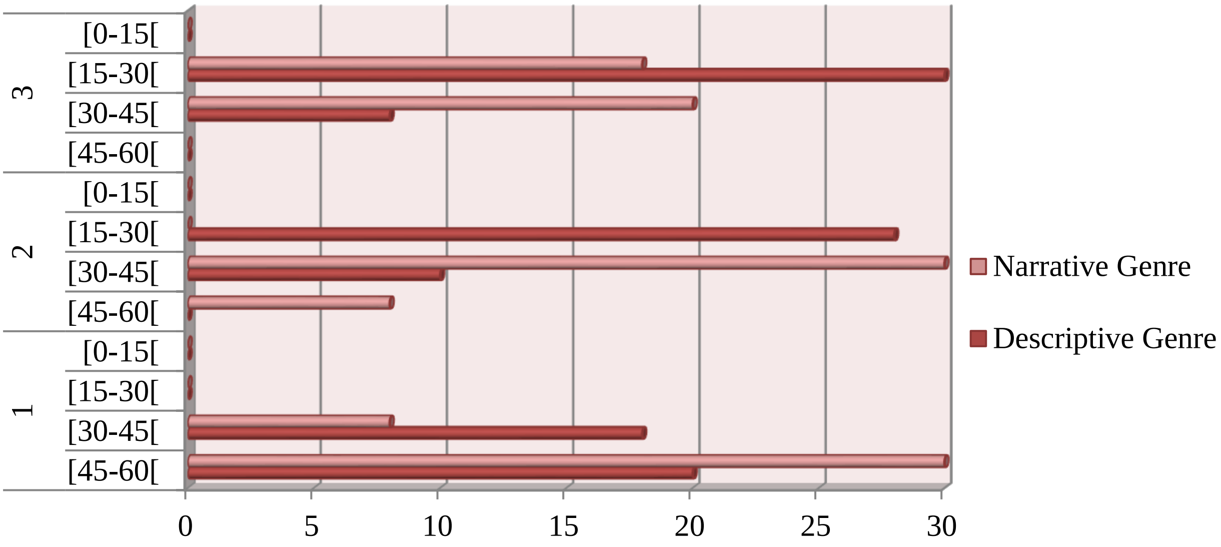
Figure 3.44 *Time Devoted to Self-assessment*



As indicated in Figure 3.44, at the beginning of the intervention all participants spent between 45 and 60 minutes to self-assess the descriptive genre, whereas the time given to self-assessing the narrative genre was usually between 15-30 minutes and 30-45 minutes. At the end of the intervention, the number of participants who devoted time to self-assess the narrative genre between 15 and 30 min. is higher as compared to self-assessing the descriptive genre in the same interval. In addition, the number of participants self-assessing the descriptive genre in lower than 15 min. is higher than the ones self-assessing the narrative genre.

3.2.1.3 Time Devoted to Anonymous Peer Assessment Performance. The researcher kept track of the time spent peer assessing both descriptive and narrative genres. Figure 3.45 depicts the outcomes**.**

Figure 3.45 *Time Devoted to Anonymous Peer Assessment*

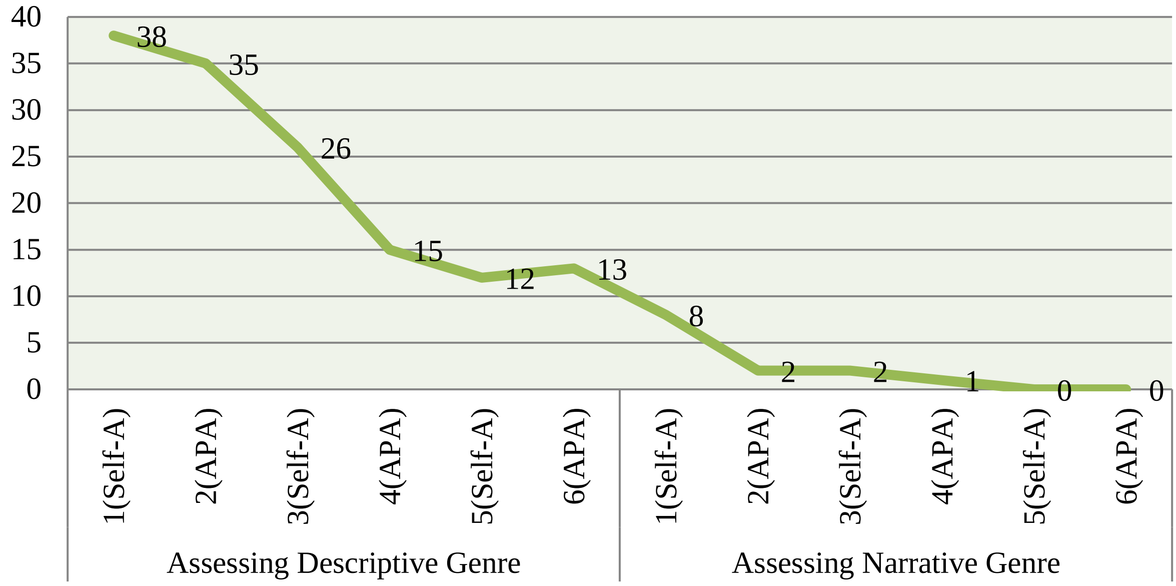


Similarly to self-assessment, as shown in Figure 3.44, the participants were quicker when peer-assessing the descriptive genre as compared to the narrative one. At the end of the intervention, the number of participants who spent between 15 and 30 min. peer assessing the narrative genre was lower than those who were peer assessing the descriptive genre. None of the participants had completed peer assessment for either genre in less than 15 min.

3.2.1.4 Individual Inquiries about Writing Rubric Use. The descriptive and narrative genres were under study. Each genre was assessed three times for each student assessment. Individual questions from participants regarding using rubrics included questions about the descriptive levels and follow-up sections in self- and APA.

a.Individual Inquiries about Descriptive levels. The researcher developed a hard copy of two pages without any descriptors but simply the name of the criteria and sub-criteria to enable participants recall the description levels for performing assessment smoothly and effortlessly. The aim was to assist participants in summarizing the descriptions in order to speed up their knowledge and memorization. This was maintained at the participant’s request while doing the assessment; if the participant needed it, they requested it. Figure 3.46 depicts the frequency of inquiring the memorization sheet.

Figure 3.46 *Inquiries' Frequency about the Descriptive Levels*

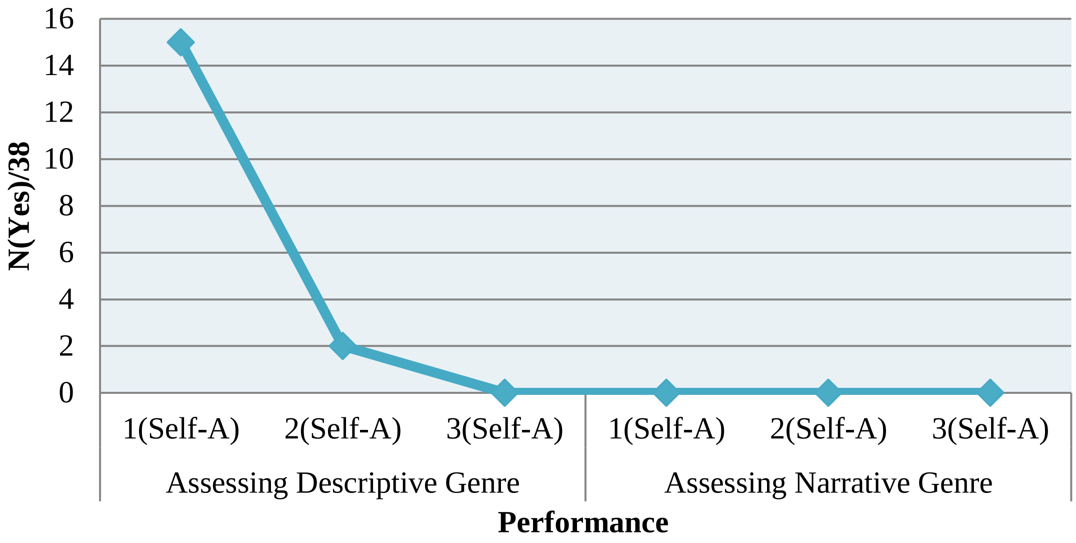


Individual inquiries regarding the descriptors decreased gradually from the first to the last performance for both genres including both student assessment performances, with a minor rise when just starting assessing the narrative genre. The frequent questions asked were about the abbreviations of the criteria and sub-criteria, being reminded about the descriptors of the criterion 'knowledge of vocabulary', and reviewing sentence problems. In addition to that, some participants asked about spelling some words, their meaning, and the translation from Arabic to English. Other questions were also asked such as the way of evaluating negation and how to judge a given controlling idea in the topic sentence to be excellent, good, average, or poor.

b.Individual Inquiries about Follow-up Sections

- **Inquiries about Follow-up Section in Self-assessment Performance.** The assessment rubrics which were associated with a follow-up section in self-assessment included two major sections: the first was about participants' dis/satisfaction with their writing and the second was about participants' dis/satisfaction of using the rubric. The frequency of inquiries related to this section is displayed in Figure 3.47.

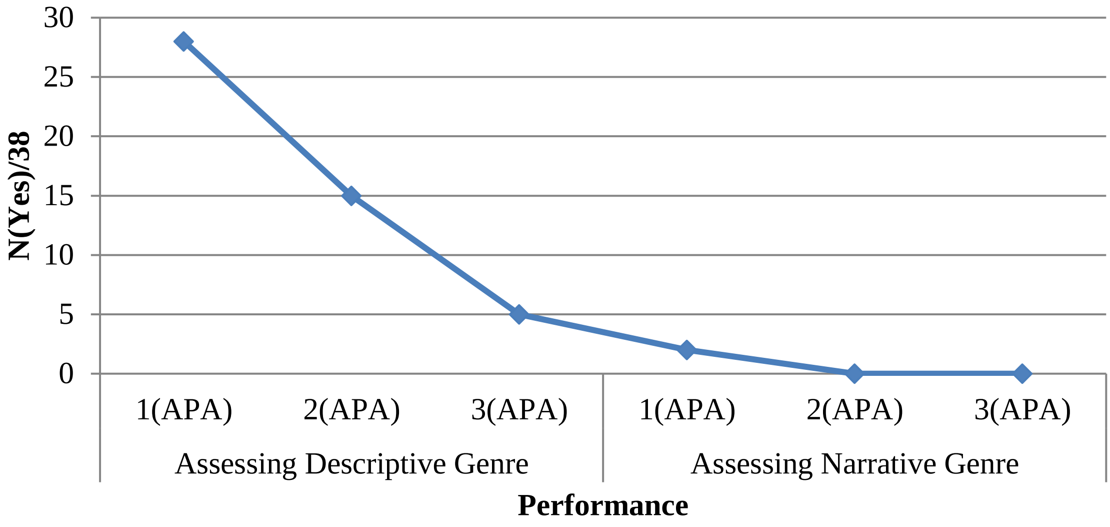
Figure 3.47 *Inquiries' Frequency about the Follow-up Section in Self-assessment Performance*



The inquiries about the follow-up section devoted to self-assessment disappeared after the third practice. The most repeated question was whether it is a must to write in all the sections or just in the corresponding section. It means whether the participant had to write in the section of satisfaction when s/he was only satisfied, or s/he had to write in all sections.

- **Inquiries about Follow-up Section in Anonymous Peer Assessment Performance.** The rubric associated with a follow-up section for APA included two major sections. The first section was about the assessors and their opinions about their classmates' writing and their opinion about the role of rubric in assessing their classmates' writings, and the second one was devoted to the assessees' responses to their classmates' assessment. The frequency of inquiries related to this section is displayed in Figure 3.48.

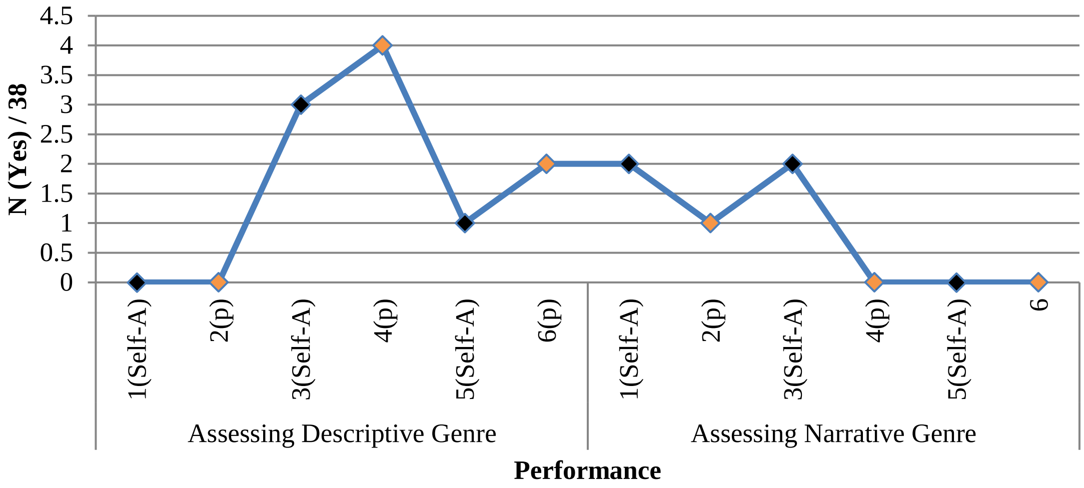
Figure 3.48 *Inquiries' Frequency about Follow-up Section in Anonymous Peer Assessment*



In contrast to the inquiries about the follow-up section in self-assessment, the ones about the follow-up section in peer assessment had gradually diminished till they disappeared. The questions asked were about how to deliver peer feedback. For example, the majority of the participants at the beginning asked about what to include in peer feedback; is it to discuss their classmates' writing weaknesses only, pointing to their strengths or both? Whether to include advice and encouragements? and so on.

3.2.1.5 Overall Classroom Assessment. The questions received regarding writing classroom assessment were mostly on the objective of this classroom assessment and its significance in respect to short and long-term goals. The frequency of these queries is seen in Figure 3.49.

Figure 3.49 *Inquiries' Frequency about Overall Classroom Assessment*



The participants' inquiries about overall writing classroom assessment started appearing after the first practice. Various questions were asked like the importance of this assessment in relation to students' writing, whether being included in the final exam, the importance of rubrics and student assessment in future, whether this classroom assessment had been used by all teachers, the importance of saying excellent, good, average, and poor without marks, the importance of marks, and so forth.

3.2.2 Analysis of Participants' Works in the Intervention Proper

Before delving into the analysis itself, three important parts are discussed in this section: the sampling technique and procedure, the assessment grid used to determine the type of assessment performed, and the quantitative assessment of the participants' levels.

Purposive sampling, which is commonly used in qualitative and mixed methods research, is adopted on purpose by the researcher to elucidate specific theme, phenomenon or concept (Robinson, 2014). Purposive sampling was used by the researcher to examine the work done by participants in the intervention itself. Six works were chosen to accomplish this goal. Each work includes all of the participant's writings and student assessments done throughout the intervention period. Participants' levels were classified as low, average, and good. Those values represented the participants’ baselines prior to the intervention. Excellent level was omitted since no participant had that level in writing at the start of the intervention. Two participants were chosen for each level. The researcher chose one female participant and one male participant. This could not be applied for good level because no female participant was found having a good writing performance level. It is worth noting that APA was performed in three practices for each genre, as follows: participant with low level associated with participant with high level and vice versa, low with low and high with high level, and lastly random assignment.

On the other hand, as there is no standard assessment grid and none found in experimental studies in the literature to go in line with the purpose of this study, the researcher established her own grid regarding quantitative assessment done (see Table 3.2). To do that, she had focused on the studies discussing under- and over-estimation concerning student assessment, self- and peer assessment. This point is fully discussed in self- and peer assessment sections in the literature review chapter (see Sections 1.7.4 & 1.7.5).

Table 3.2 *Suggested Assessment Decision Grid*

Difference		**Assessment decision**
Higher	[0-1]	Objective assessment (OA)
	]1-4]	Overestimation(OE)
	>4	High overestimation (HOE)
Lower	[0-1]	Objective assessment (OA)
	]1-4]	Underestimation (UE)
	>4	High underestimation (HUE)

To understand Table 3.2 above, if the difference between two assessments is ±1 or ≤1, that means the assessment done is objective. If the difference ranges between 1 and 4, excluding 4, and it is higher than the score given, this means that there is an overestimation. It is the same for underestimation but only if the difference is less than the given score. Finally, if the difference is higher than 4, including 4 in the interval, this means it is high overestimation in case it is higher than the given score and underestimation when it is less than the given score.

The entire assessment determines the score reference for comparative purposes. For example, if the overall assessment in self-assessment is deemed to be objective and similar to the one done by the teacher when compared to peer assessment, the assessment reference between self- and PA is self-assessment. As a result, the reader must pay attention to the column heads of the assessment decision since they do not have the same title. For example, comparing self- versus APA indicates that the reference score is the one obtained in APA.

Table 3.3 shows the intervals for judging whether a given participant has a poor, average, good, or excellent level in writing.

Table 3.3 *Quantitative/Scoring Assessment Decision*

Levels	Poor	Average	Good	Excellent
Scoring interval	[4.13- 9.25[	[9.25- 13.88[	[13.88- 16.50[	≥16.50

The major aspects focused on to analyze the content of the participant’s work were: (1) reporting the assessors’ level in peer assessment, (2) commenting on the student assessment done, both quantitative and qualitative, (3) commenting on students’ reaction to peer assessment, and (4) their behavior toward assessing their classmates’ writings.

3.2.2.1 Participants with Poor Level

Participant Number One (Female: Leila)

a. Peer Assessors’ Level. The writing level of the assessors who participated in peer assessing this participant’s written pieces is shown in Table 3.4.

Table 3.4 *Peer Assessors’ Level*

Practice(P)	Anonymous assessor level	
	Descriptive	Narrative
P1	Good	Average
P2	Poor	Poor
P3	Average	Average

b. Assessment of the Participant's Writings. As for the quantitative assessment, the participant's scores obtained from self-, peer, and TA are displayed in Table 3.5.

Table 3.5 *Quantitative Assessment and Assessment Decision*

Writing practice	Scoring			Assessment decision		
	SA	**APA**	**TA**	**Comparing APA to SA**	**Comparing SA to TA**	**Comparing APA to TA**
D1	9.13	8.38	8.00	OA	OA	OA
D2	8.88	15.63	10.88	HOE	UE	HOE
D3	10.38	11.00	11.00	OA	OA	OA
N1	8.38	10.13	11.00	OA	UE	UE
N2	9.25	14.50	13.88	HOE	HUE	OA
N3	9.25	13.88	14.00	HOE	HUE	OA

Despite being assessed by assessors having higher levels than hers, except for two, they all tended to overestimate her work. When her scores were compared to the teacher's, it was revealed that the participant tended to underestimate her work as she attempted to offer better writings. In terms of peer feedback, all of the assessors, excluding the third, tried to provide constructive criticism to help her improve her writing. For example, one of the assessors stated that she had to work on her vocabulary to gain more words. For the third assessor's feedback, the content was not correct. The fourth, the fifth and the sixth one, the feedback given was brief but targeting both the positives and the negatives in participant's writings. Also, it was written in a courteous way. For the first and the second assessors, the feedback provided was very detailed focusing on both weaknesses and strengths and was also written softly.

c. Participant's Reaction to Peer Feedback Given. The participant did not accept all of the comments that were provided. She was unpleasant and unresponsive to the criticisms leveled at her work by the first and second assessor. She did not accept the third feedback since the assessor did not offer constructive feedback and only said that her writing was nice. She justified that via announcing that feedback did not reflect the reality of her writing, as many flaws remained, such as unsuccessfully connecting topics and applying rubrics. The participant showed progressive acceptance to feedback given and being responsive to it. As an example of the first case, when she refused to accept the feedback, the participant said, ‘No, this is not my real level in writing. My writing is good because I am working hard’. In the second case, where she started being responsive to the feedback given, the participant said, ‘Yes, indeed all these remarks are true and that will help me to know my weaknesses. I will take all that into consideration’. For the third feedback, she definitely, as mentioned just above, criticized it through showing its inappropriate content as it was not targeting the pitfalls in her writing.

d. Participant’s Reaction to the Follow-up Section in Self-assessment. The participant reacted to the follow-up section accordingly. More details about that, the comments and her satisfaction or not, are summarized in Table 3.6.

Table 3.6 *Participant’s Reaction to the Follow-up Section in self-assessment*

Writing practice	Student’ s writing	Comment	Rubric use by student	Comment
D1	S	The participant was serious when dealing with her writings. When she was satisfied with her writing, she expressed clearly what satisfied her describing that in details. At the beginning, she used to evoke ideas randomly but with time she started organizing her ideas using a clear and understood	S	The same happened when using the rubric. The participant was struggling with rubric use. She perceived it easy in the first practice but with time she tried to spot the difficulties she faced. Overall, she appreciated the help of the rubric to motivate her work harder.
D2	S		S	
D3	NS		NS	
N1	NS		S	
N2	NS		S	
N3	NS		S	

		style.		

(D): Descriptive; **(N):** Narrative; **(S):** Satisfied; **(NS):** Not satisfied

The third column of Table 3.7 illustrates the presence or absence of the categories acquired in the participant's reaction to the follow-up section in contrast to the findings obtained in the post-study questionnaire. The percentages in the table's second column come from the post-study questionnaire. They are just here to remind the reader of the frequency of each category.

Table 3.7 *Participant's Reaction to the Follow-up section in Self-assessment in Comparison to the Post-study Questionnaire Results*

Category	%	Presence
Tracking progress	26	✓
Motivating section	16	✓
Communicating students' problems to the teacher	31	✓
Reminding students about their mistakes	17	✓
Disliking this section	05	✗
Repetitive section	05	✗

e. Participant and Rewriting. The participant attempted to compose longer and well-organized paragraphs than they had written at the beginning. The paragraphs in the first, second, third, and fourth rewritings were characterized by lengthy sentences with language use difficulties, but the paragraphs in the last three rewritings featured less sentence problems and became more correct.

f. Participant's Behavior toward Giving Peer Feedback. Despite her low writing level, the participant attempted to evaluate her peers' writing fairly, but in some situations she underestimated them. In terms of peer feedback, the participant tended to be straightforward, especially when offering criticism on their classmates' work for the first time: 'Your writing is bad...', 'Your paragraph is not academic...', and so on. She gradually began to include mitigators into producing her peer feedback. Initially, she used to make comments on her classmates' works, focusing solely on their flaws.

Gradually in the time, she attempted to elicit both errors and strong points in their peers' work, expressing them softly via using mitigators. As an example, one participant wrote, 'This paragraph is simple with basic ideas. In fact, there are no new ideas to attract the reader, but the time order of the narrative genre was fully respected. My dear classmate you should pay attention and you must read more to have a literacy sense in your writing'.

g. Profiling the Participant in Writing Assessment. Despite having poor writing level, this participant was eager to improve her writing. However, she kept her flaws to herself and did not share them with others. She practiced becoming an extrovert and discussing her difficulties with her peers. As a response to the results shown in Table 3.39, entitled Participants' Works Ordered, this participant's writing improved from poor to good.

Participant Number Two (Male: Zaki)

a. Peer Assessors' Level. The writing level of the assessors who participated in peer assessing this participant's written pieces is shown in Table 3.8.

Table 3.8 *Peer Assessors' Level*

Practice(P)	Anonymous assessor level	
	Descriptive	Narrative
P1	Good	Average
P2	Poor	Poor
P3	Poor	Poor

b. Assessment of the Participant's Writings. As for the quantitative assessment, the participant' scores obtained from self-, peer, and TA are displayed in Table 3.9.

Table 3.9 *Quantitative Assessment and Assessment Decision*

Writing practice	Scoring			Assessment decision		
	SA	**APA**	**TA**	**Comparing SA to APA**	**Comparing SA to TA**	**Comparing APA to TA**

D1	13.13	7.38	6.00	HOE	HOE	OE
D2	11.25	9.38	9.00	OE	OE	OA
D3	12.38	13.38	10.00	OA	OE	OE
N1	12.25	9.88	9.25	OE	OE	OA
N2	13.50	10.38	10.88	OE	OE	OA
N3	12.88	9.88	11.25	OE	OE	UE

Despite having varying degrees of writing level and largely being poor, the assessors, with the exception of the third, who inflated his classmate's writing, attempted to objectively, albeit somewhat overestimated, rate their classmate's writing. Regarding peer criticism, all of the assessors, with the exception of the one who overvalued his classmate's writing, offered highly specific comments when identifying mistakes with their classmates' writing. The way of delivering feedback was different. Excluding the third peer feedback, which was deemed incorrect and irrelevant to the content, the second and fourth assessors focused solely on their classmates' pitfalls, describing them in a rude manner, whereas the first, fifth, and final assessors focused on both sides, encouraging the writer to pay attention to his mistakes and practice more.

c. Participant's Reaction to Peer Feedback Given. The participant welcomed peer feedback from his peers, but he pointed out the unpleasant way in which some of the assessors had written it; the second and fourth ones. Except for that, he briefly indicated his acceptance of the peer feedback provided.

d. Participant's Reaction to the Follow-up Section in Self-assessment. The participant reacted to the follow-up section accordingly. More details about that, the comments and his satisfaction or not, are summarized in Table 3.10.

Table 3.10 *Participant's Reaction to the Follow-up Section in Self-assessment*

Writing practice	**Student' s Writing**	**Comment**	**Rubric use by student**	**Comment**
D1	S & NS	When it came to his writing, the participant was serious.	S & NS	When writing about rubric use, participants tended to be

D2	S & NS	He wrote briefly on what satisfied and dissatisfied him about his work. For example, he said, ‘I am not satisfied of my writing because I don’t have a lot of words’ and ‘I am satisfied because I start using correctly punctuation marks’.	S & NS	repetitious, in contrast to the writing section dis/satisfaction. He emphasized the significance of the rubric and the necessity for practice.
D3	S & NS		S & NS	
N1	S & NS		S & NS	
N2	S & NS		S & NS	
N3	S & NS		S & NS	

The third column of Table 3.11 illustrates the presence or absence of the categories acquired in the participant’s reaction to the follow-up section in contrast to the findings obtained in the post-study questionnaire. The percentages in the table’s second column come from the post-study questionnaire. They are just here to remind the reader of the frequency of each category.

Table 3.11 *Participant’s Reaction to the Follow-up Section in Self-assessment in Comparison to Post-study Questionnaire Results*

Category	%	Presence
Tracking progress	26	✓
Motivating section	16	✓
Communicating students’ problems to the teacher	31	✗
Reminding students about their mistakes	17	✗
Disliking this section	05	✗
Repetitive section	05	✓

e. Participant and Rewriting. The participant did not rewrite his paragraph for the first and second practice, but he did for the four remaining ones. Apart from the fourth practice, where the participant completely changed the idea of the paragraph, the rewritings did not see major changes. In terms of writing improvement, only the fifth and last rewriting had been acceptably improved, namely in using correct punctuation marks and appropriate language use.

f. Participant’s Behavior toward Giving Peer Feedback. Most of the feedback given by this participant to his classmates was brief including compliments. He started detailing his feedback in the sixth peer assessment, pointing out to the positive and negative aspects of his classmates’ writings.

g. Profiling the Participant’s in Writing Assessment. Despite his poor writing level, this participant made less effort than the first and showed less enthusiasm in improving his writing. As a response to the results displayed in Table 3.39, entitled Participants’ Works Ordered, this participant moved from poor to average level.

3.2.2.2. Participants with Average Level

Participant Number One (Female: Anissa)

a. Peer Assessors’ Level. The writing level of the assessors who participated in peer assessing this participant’s written pieces is shown in Table 3.12.

Table 3.12 *Peer Assessors’ Level*

Practice(P)	**Anonymous assessor level**	
	Descriptive	**Narrative**
P1	Good	Good
P2	Average	Average
P3	Average	Poor

b. Assessment of the Participant’s Writings. As for the quantitative assessment, the participant’s scores obtained from self-, peer, and TA are displayed in Table 3.13.

Table 3.13 *Quantitative Assessment and Assessment Decision*

Writing practice	Scoring			Assessment decision		
	SA	**APA**	**TA**	**Comparing APA to SA**	**Comparing SA to TA**	**Comparing APA to TA**
D1	12.5	13.25	11.00	OA	OE	OE

D2	13.88	13.50	12.88	OA	OA	OA
D3	11.50	11.88	10.88	OA	OA	OA
N1	14.38	9.88	12.25	HUE	OE	UE
N2	15.63	14.63	14.25	OA	OE	OA
N3	14.25	13.63	15.25	OA	OA	UE

The assessment was generally objective. Except for the last one, which was meaningless, peer feedback was helpful and informative, identifying both the shortcomings and positives of the written works. The fifth feedback was brief in comparison to the previous four. As an example of constructive feedback, one of the assessors stated, ‘I think that my classmate’s writing is average’. S/he noticed that the assessee had expressive ideas but not well-connected. She furthered advising the assessee to work on cohesion and coherence and punctuation marks, especially commas. S/he ended the feedback by saying, ‘your paragraph is structured and clean, and your handwriting is clear’.

c. Participant’s Reaction to Peer Feedback Given. The participant attempted to negotiate her peers’ assessment and did not readily accept them. Every time she received assessments, she pointed out the correct or incorrect usage of the rubric. Despite the praises on her work, she felt the last feedback to be ineffective since it was inaccurate. She described it as ‘useless’. She also rejected the fourth feedback since it was not delivered appropriately according to the rubric criteria.

d. Participant’s Reaction to the Follow-up Section in Self-assessment. The participant reacted to the follow-up section accordingly. More details about that, the comments and her satisfaction or not, are summarized in Table 3.14.

Table 3.14 *Participants’ Reaction to the Follow-up Section in Self-assessment*

Writing practice	**Student’ s writing**	**Comment**	**Rubric use by student**	**Comment**
D1	S	The participant stated that she was satisfied and that	S	Similarly to the follow-up section for student writing, the

D2	S	was in the first and second follow-up sections. Beginning with the third practice, the participant attempted to describe what did and not satisfy her and why, referring directly to the cause without moving back and forth. For example, she said ‘what satisfied me the most is my improvement in using the punctuation marks.’	S	participant stated that she was satisfied with the use of rubrics without expressing why. However, in the third practice, the participant emphasized the significance of rubric usage, listing the help it provided. As a result, she emphasized the progress she made with the rubric in using punctuation marks and being encouraged to use new vocabulary in her writing.
D3	NS		S	
N1	S		S	
N2	S		S	
N3	NS		/	

The third column of Table 3.15 illustrates the presence or absence of the categories acquired in the participant’s reaction to the follow-up section in contrast to the findings obtained in the post-study questionnaire. The percentages in the table’s second column come from the post-study questionnaire. They are just here to remind the reader of the frequency of each category.

Table 3.15 *Participant’s Reaction to the Follow-up Section in Self-assessment in Comparison to Post-study Questionnaire Results*

Category	%	Presence
Tracking progress	26	✓
Motivating section	16	✓
Communicating students’ problems to the teacher	31	✗
Reminding students about their mistakes	17	✗
Disliking this section	05	✗
Repetitive section	05	✗

e. Participant and Rewriting. This task had been assigned differently by the participants. There was no significant change in ideas between the first, second, and fourth rewrites, but there were significant improvements in punctuation marks for the first, writing style for the second, and spelling problems for the fourth. The paragraphs were shorter in the fifth and sixth practices and longer in the third.

f. Participant's Behavior toward Giving Peer Feedback. This participant began by providing quick feedback, but she included both positive and negative aspects of her peers' work. She was forthright during the first two practices, but after that, she began elaborating the feedback in a kind manner. For example, she said, 'I think that my classmate's writing is good, but he or she needs some effort because I found some sentences incomplete in meaning.'

g. Profiling the Participant's in Assessment and Writing. Despite her average writing level, this participant worked diligently on her writing, addressing any feedback and reacting properly to enhance her writing. As a consequence of the results shown in Table 3.39, entitled Participants' Works Ordered, this participant's writing level had improved from average to good.

Participant Number Two (Male: Ilyes)

a. Peer Assessors' Level. The writing level of the assessors who participated in peer assessing this participant's written pieces is shown in Table 3.16.

Table 3.16 *Peer Assessors' Level*

Practice (P)	Anonymous assessor level	
	Descriptive	Narrative
P1	Good	Good
P2	Average	Average
P3	Poor	Average

b. Assessment of the Participant's Writings. As for the quantitative assessment, the participant's scores obtained from self-, peer, and TA are displayed in Table 3.17.

Table 3.17 *Quantitative Assessment and Assessment Decision*

Writing Practice	Scoring			Assessment decision		
	SA	**APA**	**TA**	**Comparing SA to APA**	**Comparing SA to TA**	**Comparing APA to TA**
D1	12.13	12.13	10.00	OA	OE	OE
D2	14.38	11.00	14.00	OE	OA	UE
D3	14.00	12.68	11.75	OE	OE	OA
N1	13.75	15.25	13.88	OE	OA	OE
N2	17.13	13.5	13.88	HOE	HOE	OA
N3	15.25	12.00	14.13	OE	OE	UE

Peer assessments provided tended to overestimate the participant's writing. The participant enjoyed and expressed interest in the feedback content provided by the first, fourth, and sixth assessors since it was highly accurate, constructive, and well written. One of the assessors, for example, stated, 'I think that my classmate's writing was not clear enough because I did not get the idea. He or she did not express well his/her ideas and the use of the pronoun ''you' was unnecessarily'. The second assessor's feedback was unclear, and the participant showed no interest in it; 'I respect and appreciate the assessment of my classmate, but unfortunately she or he is not respecting the rules. In addition to that, there is a contradiction between his or her assessment using the rubric and the feedback she or he gave. I hope next time takes the writing work seriously'. For the third assessor, the feedback was only compliments and the participant pointed to not focusing just on the positive sides when assessing to make the feedback credible. For the fifth feedback, the content was brief and the participant did not show interest in it.

c. Participant's Reaction to Peer Feedback Given. The participant accepted peer feedback and pointed to the appropriate use of the rubric writing, mentioning that briefly. This was applied for the first, fourth, fifth and sixth assessors. For the second assessor, the participant commented on the feedback given pointing to the inappropriate use of the rubric and the contradiction reported between the assessment using the rubric and the peer feedback given. For the third feedback, the participant showed much importance to that saying, ‘First off, I appreciate and respect my classmate’s peer assessment of my writing. It is a good way to know whether my writing is affecting others positively or negatively. She or he respected the rules of the rubrics in most of the work and this was a good strategy to know our levels’.

d. Participant’s Reaction to the Follow-up Section in Self-assessment. The participant reacted to the follow-up section accordingly. More details about that, the comments and his satisfaction or not, are summarized in Table 3.18.

Table 3.18 *Participant’s Reaction to the Follow-up Section in Self-assessment*

Writing practice	**Student’ s writing**	**Comment**	**Rubric use by student**	**Comment**
D1	S & NS	The participant accurately and concisely explained what satisfied him in the first and second writing pieces. In the third and fourth, the participant chose to provide more details, however in the last two, the participant just expressed his satisfaction with his writing growth.	S	The participant was satisfied with the rubric and liked the assistance it provided for each of the six practices since it motivated him to perform better.
D2	S & NS		S	
D3	S & NS		S	
N1	S & NS		S	
N2	S & NS		S	
N3	S & NS		S	

The third column of Table 3.19 illustrates the presence or absence of the categories acquired in the participant’s reaction to the follow-up section in contrast to the findings obtained in the post-study questionnaire. The percentages in the table’s second column come from the post-study questionnaire. They are just here to remind the reader of the frequency of each category.

Table 3.19 *Participant's Reaction to Follow-up Section in Comparison to Post-study Questionnaire Results*

Category	%	Presence
Tracking progress	26	✓
Motivating section	16	✓
Communicating students' problems to the teacher	31	✓
Reminding students about their mistakes	17	✓
Disliking this section	05	✗
Repetitive section	05	✗

e. Participant and Rewriting. All of the participant's writings were rewritten. The first and fourth were both lengthened, with the first receiving far more development and the fourth receiving far less. The second rewriting was not improved and appears to have been composed hastily because no punctuation marks were used in comparison to the original version. The participant struggled with punctuation marks in the first three papers, but this was gradually corrected.

f. Participant's Behavior toward Giving Peer Feedback. This participant provided brief feedback targeting only the weaknesses in his classmates' writing. Sometimes, he specified which levels, poor, average, good, or excellent, his classmates have in writing.

g. Profiling Participant's in Writing Assessment. The key changes made were largely developing better organized paragraphs in terms of form and substance, as well as improvements in employing capitalization and punctuation marks effectively and correctly. He did not learn a lot of new words and tried to produce correct structures. He began with an above-average level and finished with a low good level. As a consequence of the results shown in Table 3.39, entitled Participants' Works Ordered, one could state that this participant's writing improved from average to good.

3.2.2.3 Participants with Good Level

Participant Number One (Male: Sami)

a. Peer Assessors' Level. The writing level of the assessors who participated in peer assessing this participant's written pieces is shown in Table 3.20.

Table 3.20 *Peer Assessors' Level*

Practice (P)	**Anonymous assessor level**	
	Descriptive	**Narrative**
P1	Good	Good
P2	Good	Good
P3	Average	Good

b. Assessment of the Participant's Writings. As for the quantitative assessment, the participant' scores obtained from self-, peer, and TA are displayed in Table

Table 3.21 *Quantitative Assessment and Assessment Decision*

Writing practice	Scoring			Assessment decision		
	SA	**APA**	**TA**	**Comparing APA to SA**	**Comparing SA to TA**	**Comparing APA to TA**
D1	15.00	15.5	14.75	OA	OA	OA
D2	16.38	14.25	14.75	UE	OE	OA
D3	15.88	12.00	16.00	UE	OA	OA
N1	15.00	12.00	14.88	UE	OA	UE
N2	16.38	13.13	16.13	HUE	OA	UE
N3	16.13	15.00	17.00	UE	OA	UE

The assessors frequently undervalued their peers' writing. With the exception of the fourth and last one, the feedback they received was not specific.

c. Participant's Reaction to Peer Feedback Given. Except for the second one, where he discussed the comments made, the participant did not reply fully to the criticism offered. He thanked the assessors for their feedback each time, and he was most interested in the fourth and sixth feedback since it was so detailed and accurate.

d. Participant's Reaction to the Follow-up Section in Self-assessment. The participant reacted to the follow-up section accordingly. More details about that, the comments and his satisfaction or not, are summarized in Table 3.22.

Table 3.22 *Participant's Reaction to the Follow-up Section in Self-assessment*

Writing practice	Student's writing	Comment	Rubric use by student	Comment
D1	S & NS	When he wrote in the follow-up space, the participant was quite brief. He responded to both parts, expressing dis/satisfaction, by pointing to the important concerns simply and directly each time. In only the first and third practices, he had pointed out the flaws and strengths in his writing and topic selection.	S & NS	For the section on rubric use, the answers were the same as in the section on writing dis/satisfaction. The assessor acknowledged the value of the rubric in the area about rubric use satisfaction, and in the section about dissatisfaction he stated his desire to rate his work or to be rated each time.
D2	S & NS		S & NS	
D3	S & NS		S & NS	
N1	S & NS		S & NS	
N2	S & NS		S & NS	
N3	S & NS		S & NS	

The third column of Table 3.23 illustrates the presence or absence of the categories acquired in the participant's reaction to the follow-up section in contrast to the findings obtained in the post-study questionnaire. The percentages in the table's second column come from the post-study questionnaire. They are just here to remind the reader of the frequency of each category.

Table 3.23 *Participant's Reaction to the Follow-up Section in Self-assessment in Comparison to Post-study Questionnaire Results*

Category	%	Presence
Tracking progress	26	✓
Motivating section	16	✓
Communicating students' problems to the teacher	31	✗
Reminding students about their mistakes	17	✓
Disliking this section	05	✗
Repetitive section	05	✓

e. Participant and Rewriting. Excluding the first rewrite, the participant attempted to make appropriate adjustments while maintaining the same paragraph length, except for the fourth writing, which was a little reduced in length due to its extended length. The paragraph writer effectively improved his usage of punctuation marks and employed more complicated structures. He also used new words.

f. Participant's Behavior toward Giving Peer Feedback. This participant had provided constructive feedback while paying attention to all of the details in his classmate's work. He put short feedback when required and an extended one when necessary. He tended to undervalue his classmates' works in the first and second feedback.

g. Profiling Participant's in Writing Assessment. The participant improved in terms of a number of aspects. Despite the usefulness of several comments made by his classmates, there was a significant degree of resistance and overestimation of his writings; he worked harder and objectively assessed his writings, but he was less open to critics and feedback. As a consequence of the results shown in Table 3.39, entitled Participants' Works Ordered, this participant's writing improved from good to excellent.

Participant Number Two (Male: Adam)

a. Peer Assessors' Level. The writing level of the assessors who participated in peer assessing this participant's written pieces is shown in Table 3.24.

Table 3.24 *Peer Assessors' Level*

Practice (P)	**Anonymous assessor level**	
	Descriptive	**Narrative**
P1	Good	Good
P2	Good	Good
P3	Poor	Average

b. Assessment of the Participant's Writings. As for the quantitative assessment, the participant' scores obtained from self-, peer, and TA are displayed in Table 3.25.

Table 3.25 *Quantitative Assessment and Assessment Decision*

Writing practice	Scoring			Assessment decision		
	SA	**APA**	**TA**	**Comparing APA to SA**	**Comparing SA to TA**	**Comparing APA to TA**
D1	14.00	14.75	14.00	OA	OA	OA
D2	16.13	13.75	14.75	UE	OE	OA
D3	15.00	15.38	14.75	OA	OA	OA
N1	14.88	14.50	14.13	OA	OA	OA
N2	15.50	17.88	14.25	OE	OE	OE
N3	16.38	14.63	16.88	OE	OA	UE

The assessors had tried to objectively assess their classmates' writings. In terms of feedback, the first, fourth, fifth, and sixth assessors concentrated on appreciating his writing. However, the content was detailed in the second and third feedback. The second assessor had written discourteous feedback.

c. Participant's Reaction to Peer Feedback Given. The participant liked and enjoyed reading the feedback given by the assessors complementing on his writing, but he responded aggressively to the feedback given by the second and the third assessor. He even criticized their levels and said that they were unable to assess.

d. Participant’s Reaction to the Follow-up Section in Self-assessment. The participant reacted to the follow-up section accordingly. More details about that, the comments and her satisfaction or not, are summarized in Table 3.26.

Table 3.26 *Participant’s Reaction to the Follow-up Section in Self-assessment*

Writing practice	**Student’ s writing**	**Comment**	**Rubric use by student**	**Comment**
D1	S & NS	Except for the third practice where he detailed the two sections about writing dis/satisfaction pointing to all his strengths and weaknesses, in the remaining practices the participant selected either what satisfied or not, and he expressed that conveying that directly. In the opposite section, he wrote ‘I don’t have anything to say’.	S & NS	The participant treated this section exactly the same as he did in the one about writing dis/satisfaction. He detailed only the two sections of the third practice. However, for the other practices the participant selected just one of the two sections, either satisfaction of using rubric or dissatisfaction of using rubrics, and he wrote it briefly.
D2	S & NS		S & NS	
D3	S & NS		S & NS	
N1	S & NS		S & NS	
N2	S & NS		S & NS	
N3	S & NS		S & NS	

The third column of Table 3.27 illustrates the presence or absence of the categories acquired in the participant’s reaction to the follow-up section in contrast to the findings obtained in the post-study questionnaire. The percentages in the table’s second column come from the post-study questionnaire. They are just here to remind the reader of the frequency of each category.

Table 3.27 *Participant’s Reaction to Follow-up Section in Self-assessment in Comparison to Post-study Questionnaire Results*

Category	%	Presence
Tracking progress	26	✓

Motivating section	16	✓
Communicating students' problems to the teacher	31	✗
Reminding students about their mistakes	17	✓
Disliking this section	05	✗
Repetitive section	05	✓

e. Participant and Rewriting. This participant made no effort to rewrite his early versions. He even rewrote several of them, resulting in lower-quality versions than the originals. He just rewrote the second draft, producing an improved version by lengthening the paragraph with suitable punctuation and general language use. However, the versions created for the remaining rewrites appear to have been written carelessly.

f. Participant's Behavior toward Giving Peer Feedback. This participant's feedback to his peers was quite thorough, with underestimation at all times. For example, he assessed 'good' writing as 'average' and 'excellent' writing as 'good', and so on. Although there is some negative exaggeration, he nonetheless concentrated on all parts of an academic written piece. Despite this, he made every effort to point out all of the risks and provide guidance to his peer.

g. Profiling Participant's in Writing Assessment. This participant showed a kind of resistance toward his/her peers' feedback and tended to overestimate himself. He made no notable improvement except for acquiring some new vocabulary and writing some complicated structures. He advanced to low excellent level. As a response to the results displayed in table 3.39, entitled Participants' Works Ordered, this participant moved to excellent level in writing.

3.3 Data Analysis and Presentation: Third Stage

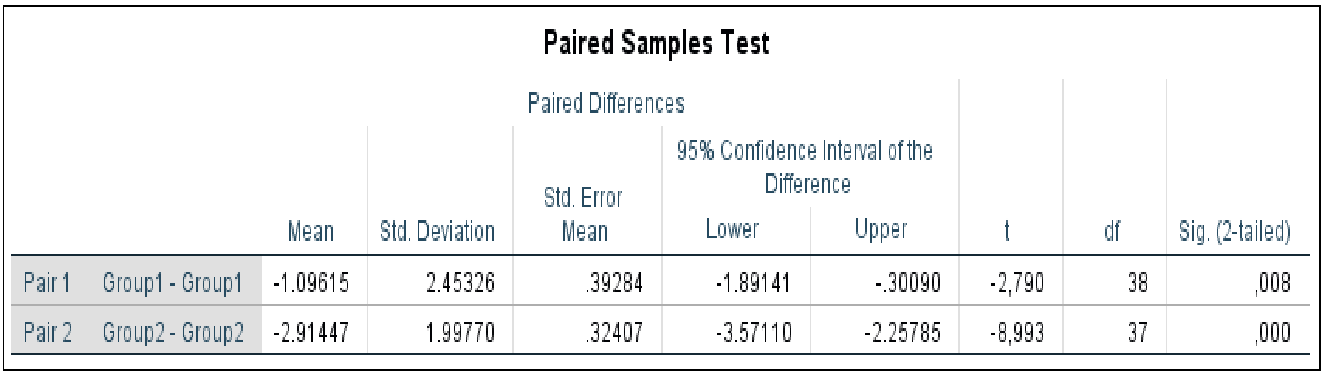
The data analyzed in this section were obtained from pre-and post-tests score analysis, post-study questionnaires, and writing working portfolio. The analysis methods used are presented in Appendix 14.b.

3.3.1 T Test Analysis

Tables 3.28 and 3.29 below were obtained through running analysis via the SPSS version 26. For the same group, paired sample T test was run and the independent sample T test was run for the control and experiment group. The P value or α is <0.05 and the critical interval is 95%. The scores are treated as interval data for analysis because the parametric tests required so (Hinton, Brownlow, McMurray, & Cozens, 2004, p.97).

3.3.1.1 Parametric Paired Sample T Test. Paired sample T tests were used to examine the differences in the means of the pre-test and post-test scores of the control and experiment groups. Table 3.28 displays the results obtained.

Table 3.28 *Results of the Parametric Paired Sample T test*



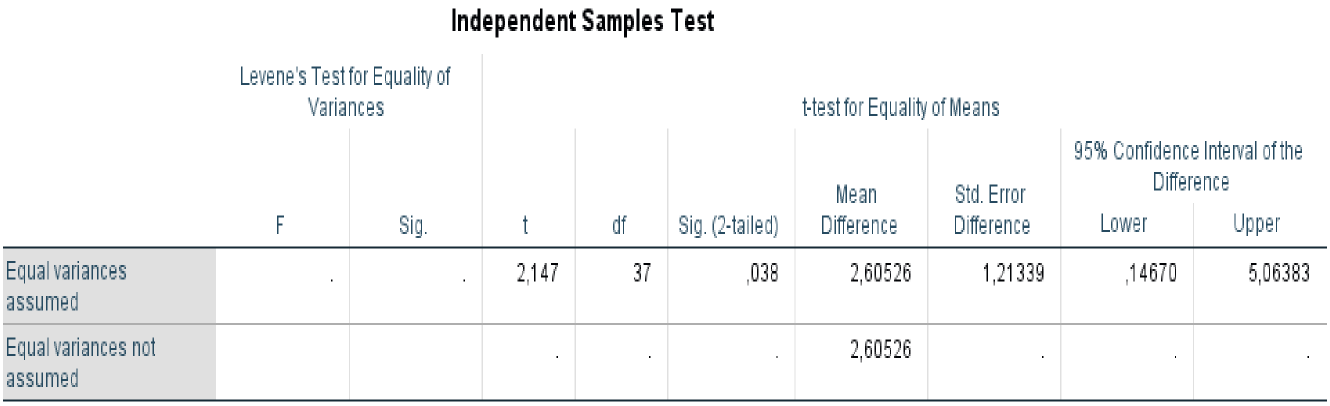
Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper			
Pair 1	Group1 - Group1	-1.09615	2.45326	.39284	-1.89141	-.30090	-2,790	38	,008
Pair 2	Group2 - Group2	-2.91447	1.99770	.32407	-3.57110	-2.25785	-8,993	37	,000

The control group is Group 1, while the experiment group is Group 2. The precise significance value for the experiment group is .00076172. Both groups, control and experimental, reveal a substantial difference, although the significance of the experiment group is higher. The difference in significance is .00724, indicating that the significance for the experiment group is nearly three times the one of the control group.

3.3.1.2 Parametric Independent Sample T Test. An independent sample T test was run to compare the means of the post-test scores of the control and experiment groups.

Table 3.29 *Results of the Parametric Independent Sample T Test*



Independent Samples Test

	Levene's Test for Equality of Variances		t-test for Equality of Means						
								95% Confidence Interval of the Difference	
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Equal variances assumed	.	.	2,147	37	,038	2,60526	1,21339	,14670	5,06383
Equal variances not assumed			.	.	.	2,60526	.	.	.

As none of the numbers of the 95% confidence interval crossed 0, the difference is seen significant in this study (Griffith, 2007). The exact value of the significance is .038.

3.3.1.3 Interpreting T test Analysis. Interpreting findings by accepting or rejecting alternative hypotheses has recently been frowned upon. Many scholars in this area do not accept statistical tests and mathematical processes such as confidence intervals (Earp & Trafimow, 2015); yet, such prohibitions have not been lifted without debate (Ashworth, 2015; Greenland, Rothman, Carlin, Poole, Goodman, & Altman, 2016). As a result, the researcher in this study did not place a great value on interpreting outcomes based on such statistical and numerical data, but rather chose to report and analyze the intervention occurrence on a regular basis while taking all aspects into account. Briefly, Based on the quantitative analysis, the result show that the confidence does not contain zero and $p \leq .05$ (0.038), which indicates that there is a significance between the pre- and post-test. On the other hand, interpreting results without relying on quantitative side, Ioannidis (2005) and Cumming (2013) advocate a strong emphasis on the method of doing research rather than merely relying on statistical and numerical results to ensure public access to the outcomes for future replications. For recall, the process of doing this study is fully described in the research method and design chapter. The T test run indicates a significant difference in the findings, and some outliers in the data are detected; however, this did not impact the overall conclusion.

3.3.2 Post-study Questionnaire Analysis

The post-study questionnaire was separated into two sections, each of which was administered in a separate session. The post-study questionnaire was divided into two parts and each was administered in an independent session. The note-taking technique was used to expound on the responses provided in the first part of the questionnaire, namely items 9 and 10 in the first section, when questioning participants about the influence of the scoring rubric in connection to the objectivity of self- and PA.

3.3.2.1 Analysis of the First Part of the Post-study Questionnaire

a. Analysis of the First Section: Rubrics' Presentation and Use

Question 1. Without exception, all participants found the rubrics' form to be simple and easy to use.

Question 2. Without exception, all participants considered the structure of the rubrics suitable in their organization.

Question 3. In terms of content, participants regarded the rubrics interesting (64%) and effective (36%).

Question 4. All participants found the descriptors easy to use because they were clear and understandable.

Question 5. The majority of participants (68%) considered the rubric function useful and helpful, while a small percentage (32%) found it extremely useful and helpful.

Question 6. When asked whether they had used the rubrics easily, all participants replied 'Yes' but after some uses: 2 uses (36%), 3 uses (29%), 4 uses (21%), 5 uses (11%). No answer for 3 % of the participants.

Question 7. All participants liked the idea of using rubrics for many reasons presented in Table 3.30. Their functions are summarized in four verbs: help, facilitate, define, and teach i.e. rubrics help, define, teach and facilitate. Table 3.30 provides further details.

Table 3.30 *Importance of Using Rubrics*

Theme	Category	%
Help	Assessing writing	28
	Spotting mistakes	18
	Improving writing	16
	Assisting both teacher and students	02
Facilitate	Assessment	06
Define	Rules to write effectively	06
	Rules to assess objectively	07
	The quality of writing	06
Teach	The rules of assessment	05
	The rules of writing	06

Question 8. The majority of participants (87%) accepted using the rubric if given another chance in the future while a few (17%) did not. For the students who responded ‘No’, the reasons were (1) difficulties to practice with the rubric (20%), (2) rubric effectiveness only when used the first time (20%), and (3) the possibility of doing assessment without rubrics (60%). However, for the participants who responded ‘Yes’, they justified their answer by various reasons, as displayed in Table 3.31

Table 3.31 *Reasons of Using Rubrics if Given a Second Chance*

Category	%
Giving a chance to assess	06
Helping in correcting mistakes	27
Helping in improving writing	06
Communicating the norms of writing	07
Helping in avoiding mistakes	10
Raising awareness	10
Criticizing writing	10
Helping in self-assessment	07
Helping in peer assessment	07
Motivation purpose	10

Question 9. The scoring rubric was appealing to all participants for a variety of reasons, which are listed in Table 3.32.

Table 3.32 *Reasons for Liking Scoring Rubric*

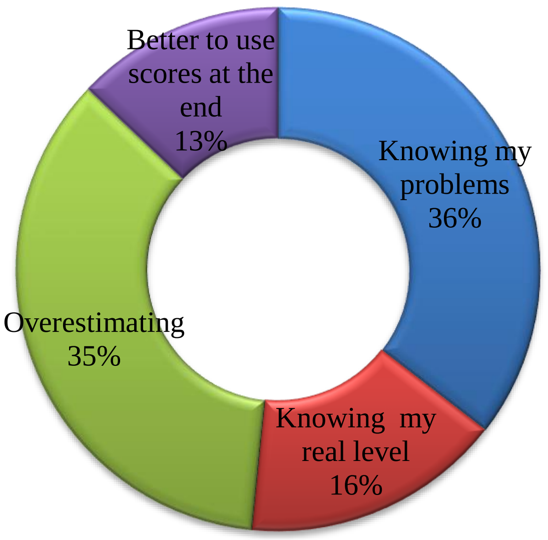
Category	Sub-category	%
Motivating toward writing progress	*Encouraging*	18
	Stressing	07
Completing the qualitative rubric		09
Reflecting student level		18
Satisfying student curiosity		07
Scoring better than remarks		18
Attractive effect of scores		05
Providing fair and objective assessment		16
Important but not necessary		02

However, some participants underlined that they should not be used right from the beginning. In this context, here are some of their answers:

1. ‘If I knew that I score my writing using the scoring rubric, I would give as much as I could good or excellent for my writings’
2. ‘It is a good idea to keep the scoring rubric at the end because that would definitely affect my qualitative assessment’
3. ‘Students are attracted by marks too much, so using this rubric from the beginning will affect assessment.’

Question 10. When questioned about the influence of employing the scoring rubric in self-assessment from the start, half of the participants (50%) believed their assessment would be subjective, while the other (44%) believed it would be objective. A small percentage (6%), only, predicted a mix of objectivity and subjectivity. Figure 3.50 depicts the reasons for and against being objective.

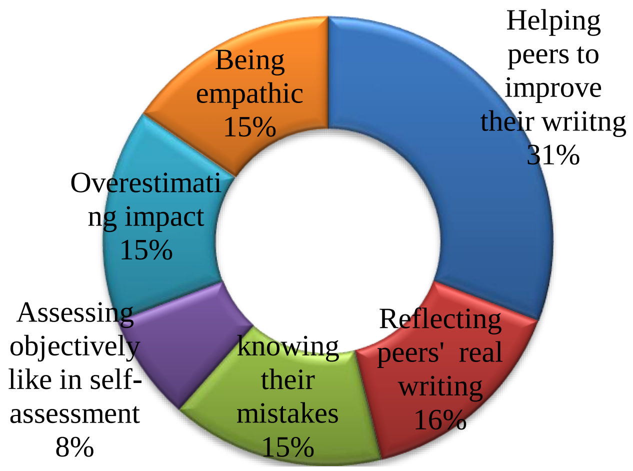
Figure 3.50 *Opinions about Using the Scoring Rubric Right from the Beginning in Self-assessment*



'Knowing my problems' and 'knowing my real level' reflect objectivity in assessment, whereas the two other categories justify the subjectivity.

Question 11. When asked about the impact of using the scoring rubric in peer assessment right from the beginning, more than half of the participants (61%) found that their assessment would be objective, others thought it would not be objective (33%) while a minority responded that it would be a mixture of objectivity and subjectivity (6%).

Figure 3.51 *Opinions about Using the Scoring Rubric Right from the Beginning in Peer Assessment*



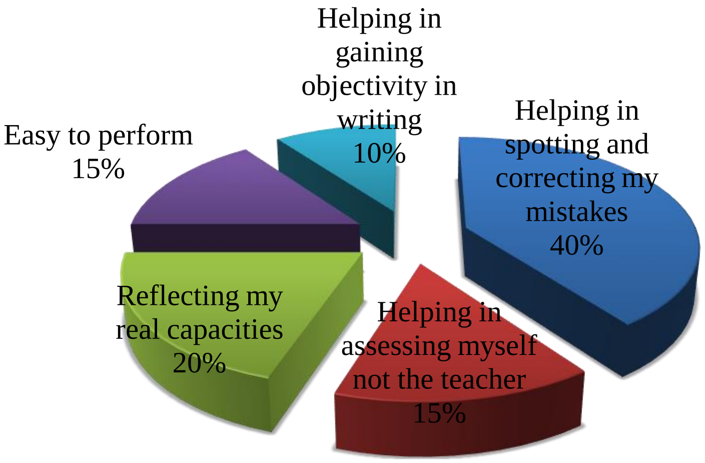
Those who believed it would not be objective were justified by: ‘being empathic’ toward and ‘overestimating’ their peers. Those who felt it would be objective were justified by providing assessments that enhance their classmates’ writing and help them understand their true level of writing and their flaws. They also expected it to be objective, as in self-assessment.

b. Analysis of the Second Section: Self-assessment Performance

Question 1. All participants, except for two, liked self-assessment performance.

Question 2. Figure 3.52 reflects how participants found self-assessment.

Figure 3.52 *Participants' Perception of Self-assessment*



Question 3. 67 participants agreed that self-assessment could not take place without rubrics. They justified their answers as follows: rubrics (a) clarify the criteria of assessment (45%), (b) preserve (protect/maintain) objectivity in assessment (33%), (c) lead to consistent assessment (11%), and (d) confer preciseness in assessment (11%). However for the participants who responded ‘Yes’, they confirmed that self-assessment could take place but (a) will be hard (29%), (b) will not be objective (29%), (c) will not be professional (14%), (d) will take more time (14%), and (e) will be incomplete assessment (14%).

Question 4. Aside from being disliked and/or regarded as repetitious by some participants, 28 participants considered the follow-up section of the rubric in self-assessment beneficial in a variety of ways. Table 3.33 detailed the arguments they presented, including example(s) for each category.

Table 3.33 *Participants' Opinions about Rubrics' Follow-up Section in Self-assessment*

Category	Example	%
Tracking progress	'This section helped me in tracking my progress by reporting each time what satisfied me and not in my writing and the use of the rubric'	26
Motivating section	'Each time I come to write in this section, I feel that I achieved a part in my writing and there is a missing part. It really motivated me'	16
Communicating students' problems to the teacher	'This section was very helpful to teacher to see her students' mistakes and problems'	31
Reminding students about their mistakes	'This section played the role of the reminder as each time says to me have you solved your problems or not?'	17
Disliking this section	'To be honest, I disliked this section because it forced me to write and I didn't have what to say'	05
Repetitive section	'I think that I was repeating each time the same thing. It is useless'	05

Question 4. A high percentage of participants (69%) indicated that they had performed objective self-assessment, while some (22%) confessed that the self-assessment they had done was extremely objective. Only a small percentage of participants (9%) answered that they had done it subjectively.

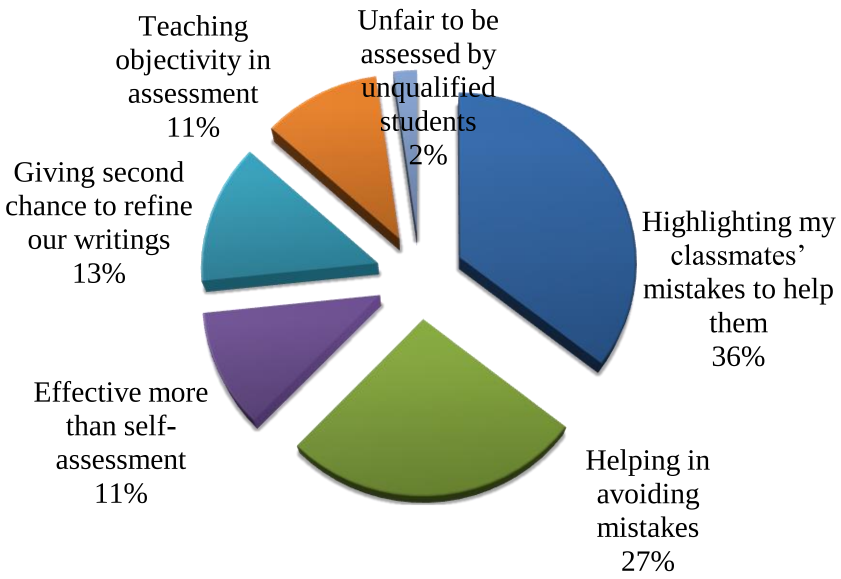
Question.5. All students accepted to undertake self-assessment if given the opportunity in the future.

c. Analysis of the Third Section: Anonymous Peer Assessment Performance

Question 1. All participants without exception welcomed APA.

Question 2. They all thought it was interesting. As a result, some of them (2%) regretted being evaluated by unqualified peers. Further clarifications are provided in discussion chapter section 4.1.4.

Figure 3.53 *Participants' Perception of Anonymous Peer Assessment*



Question 3. All participants said that APA could not take place without rubrics because the latter (a) provide detailed criteria of assessment (21%), (b) are necessary like in self-assessment (8%), and (c) help to stay objective in assessment (29%). Without rubrics, assessment can be random (19%), confused (14%), or useless (9%).

Question 4. Similar to the exceptions observed in the follow-up section of self-assessment, this section in APA was not also liked and found repetitive for some participants. As it was not alike the follow-up section in self-assessment, some participants thought that it would be trouble-making; creating conflicts between students when giving feedback to their peers.

Table 3.34 *Participants' Opinion about Rubrics' Follow-up Section in Anonymous Peer Assessment*

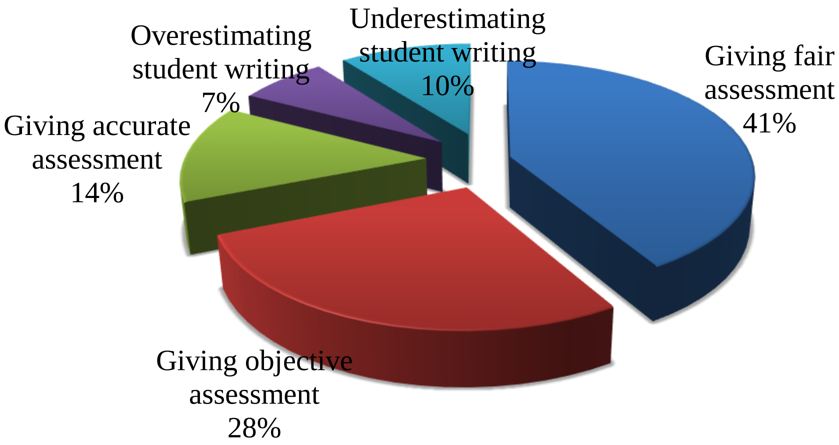
Category	Example	%
Providing constructive peer feedback	'It is really helpful when we know others' thoughts and remarks if they are honest and objective'	42

Motivating section	'This section encouraged me to reconsider my writings and look at my mistakes carefully'	10
Opening space for criticism	'It is very helpful by exchanging criticism and that would help both parties to improve their writing'	15
Reminding students about their mistakes	'Seeing my classmate's mistakes alert me not to do like that'	23
Disliking it	'I didn't like this section because it was not helpful for me'	04
Repetitive section	'I found myself repeating the same thing each time'	04
Possibility of creating conflicts	'This section could create troubles between students'	02

Question 5. Anonymous assessment was completed very objectively (46%), objectively (42%), and subjectively (12%).

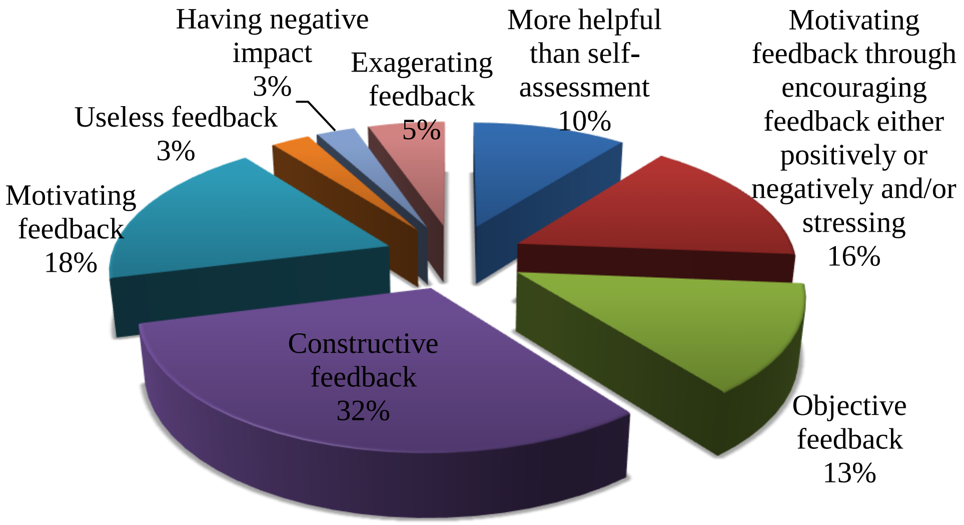
Question 6. The assessment provided by participants to their counterparts was perceived fair (41%), objective (28%), and accurate (14%). In addition to that, other participants considered it to be under- and over-estimating.

Figure 3.54 *Perception of Provided Peer Assessment*



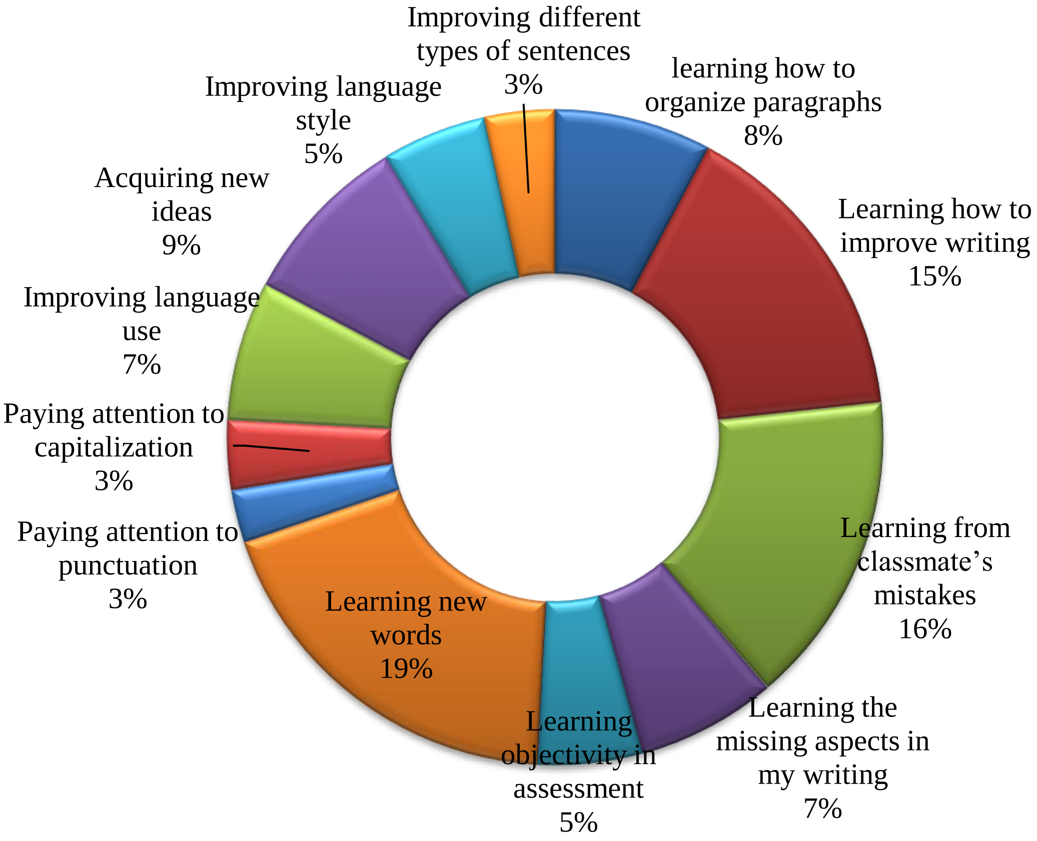
Question 7. With the exception of being viewed as exaggerated (5%), negative (3%), and worthless (3%), the majority of peer feedback was judged to be useful and fruitful in a variety of ways. Figure 3.55 depicts how participants perceived their peer feedback.

Figure 3.55 *Perception of ¨Provided Peer Feedback*



Question 8. The majority of participants (87%) stated that they learned from their peers' writings while others did not (13%). Figure 3.56 presents in details what the participants acquired from their classmates' writings.

Figure 3.56 *Peers Writings' Impact*



Question 9. All participants accepted to do APA if they were given a chance for that in the future.

3.3.2.2 Analysis of the Second Part of the Post-study Questionnaire

a. Analysis of the First section: Classroom Assessment

Question 1. Except for one, all participants preferred either APA (50%) or self-assessment (47%), but not both. They each rationalized their responses differently.

Figure 3.57 *Participants' Preference for Anonymous Peer Assessment*

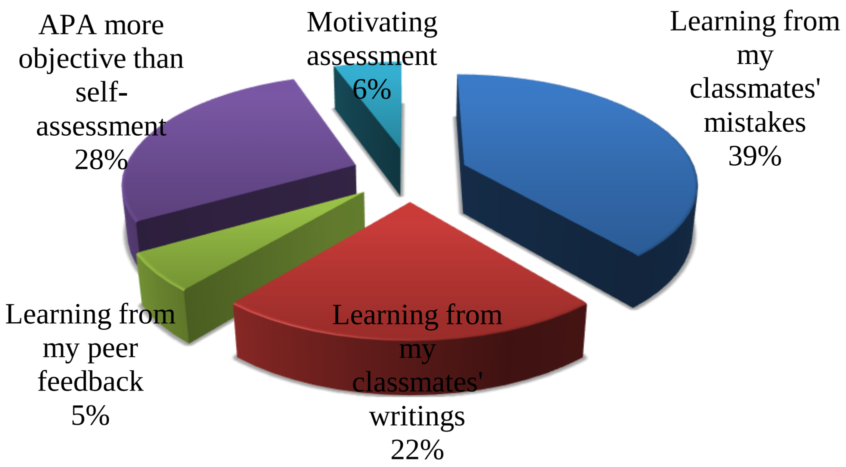
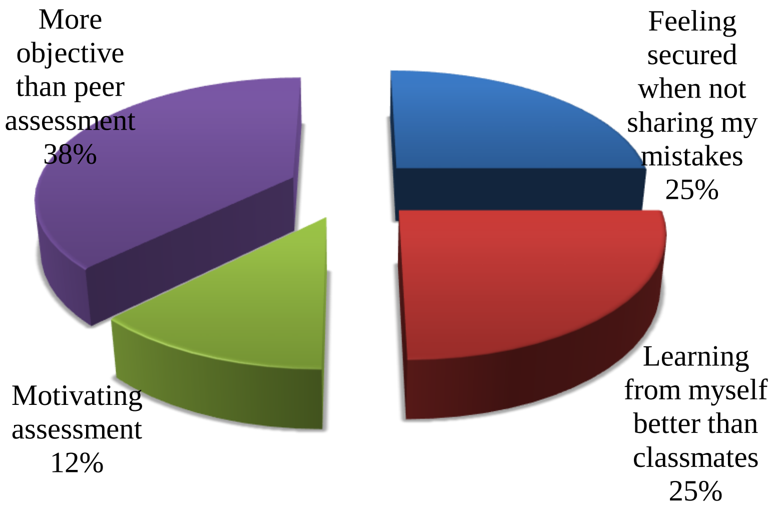
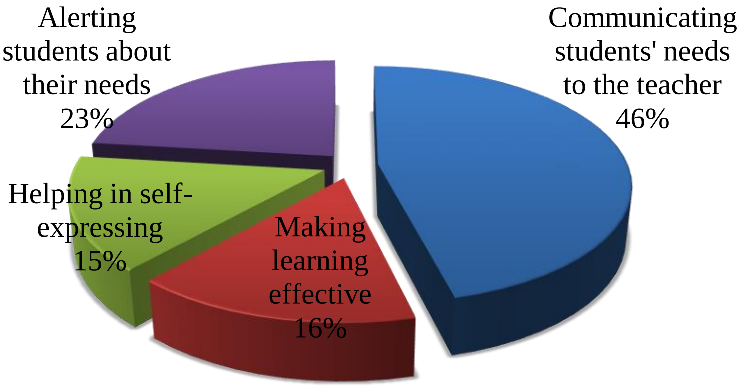


Figure 3.58 *Participants' Preference for Self-assessment*



Question 2. The majority of participants (82%) preferred providing feedback before doing any assessment. They defended their responses, as shown in Figure 3.59 below.

Figure 3.59 *Reasons of Involving Participants in Giving Opinions before Doing any Assessment*



Question 3. The majority of participants (88%) thought the time allotted was adequate for practicing writing in both genres, while a minority (6%) thought it was only adequate for the narrative genre. However, only (3%) thought time allotted was adequate for the descriptive genre, and (3%) thought it was insufficient for both.

Question 4. Without exception, all participants stated that they had used their teacher's feedback to improve their classroom assessment practice.

Question 5. When students were asked about their experience in taking part and participating in classroom assessment, their answers were principally divided into five major themes. Each had a set of categories. These details are displayed in Table 3.35.

Table 3.35 *Participants' Perception of Writing Classroom Assessment*

Theme	Category	%
Overall classroom assessment	Interesting experience	14
	Requiring much efforts	05
	Helpful most of the time	03
	Making writing less difficult	02
	Need of more practice sessions	05
	Helpful to achieve in other modules	03
	Pleasing experience	02
	Boring experience at the beginning	02

	Hard experience at the beginning	01
Student assessment performance	Important and helpful self-assessment	03
	Important and helpful peer assessment	02
	Mixing self- and peer assessment more helpful than either alone	01
Rubrics’ perception	Helpful use of rubrics	03
	Requiring more practice sessions to use rubrics	01
Writing performance	Marking progress in my writing	10
	Recognizing my mistakes and correcting them	08
	Changing my vision toward writing	03
	Feeling more professional in my writing	02
	Gaining more objectivity in assessing my writing	02
	Learning to think before writing	01
Individual traits as learners	Feeling important in the classroom	07
	Having the potential to progress	02
	Feeling motivated	01
	Helping in discovering my abilities	02
	Feeling proud of myself	01
	Feeling self-responsible	01
	Gaining self-confidence	02
	Gaining awareness	05
	Feeling scared at the beginning	01

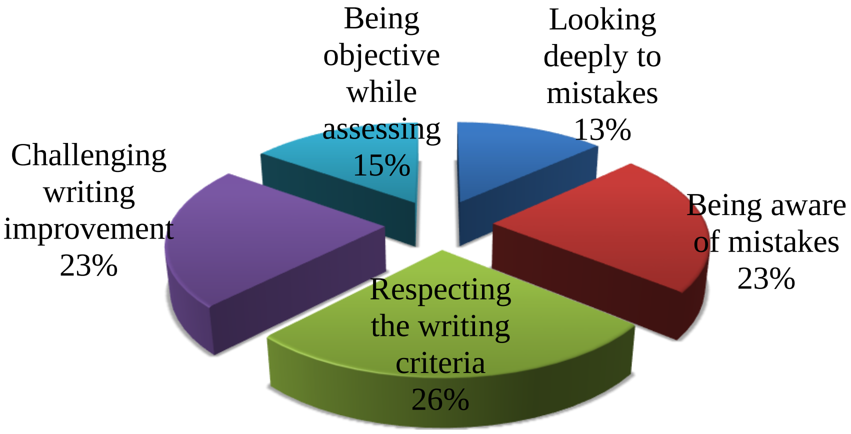
b. Analysis of the Second Section: Student Writing Progress

Question 1. All participants, except for two, felt that their writings had improved.

Question 1.1. The progress was average for both genres (59%), quick for both (17%), slow for both (9%), quick for the descriptive genre and slow for the narrative genre (6%), slow for the descriptive genre and quick for the narrative genre (3%), slow for the descriptive genre and average for the narrative genre (3%), and slow for the descriptive genre and quick for the narrative genre (3%).

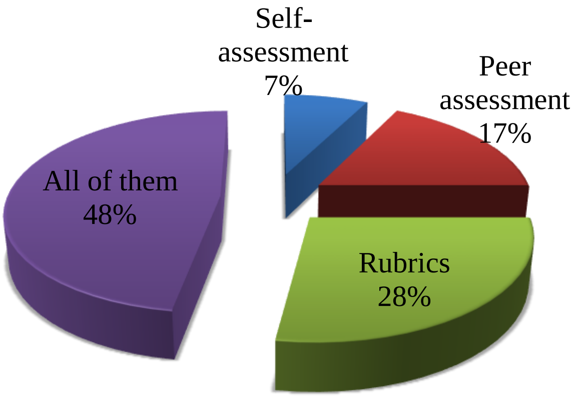
Question 1.2. When asked how they made improvement, the participants gave various explanations, as seen in Figure 3.60.

Figure 3.60 *Ways of Achieving Progress in Writing*



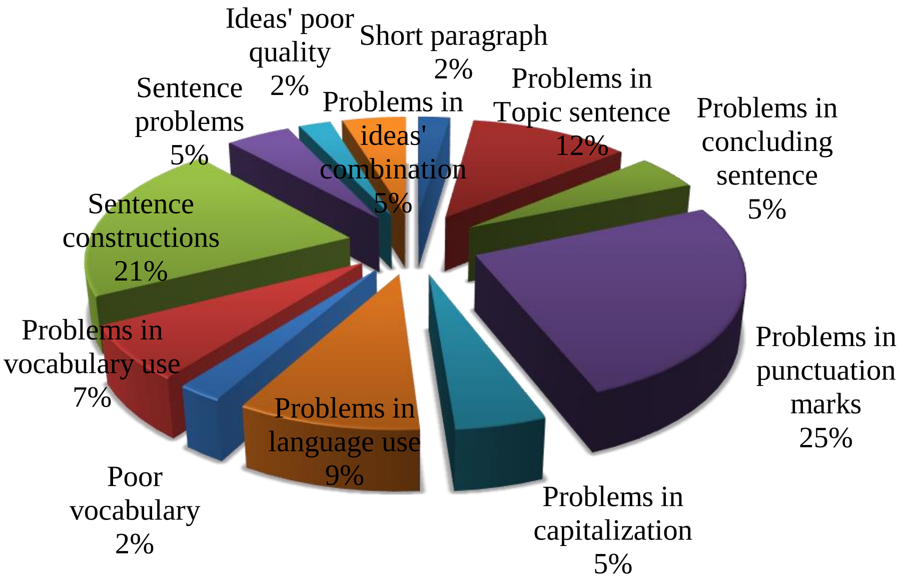
Question 1.3. This progress was mostly attributed to the combination (48%) of rubrics, self- and peer assessments.

Figure 3.61 *Ways of Helping Participants to Progress*



Question 2. The majority of participants (77%) admitted still having weaknesses. These problems are presented in Figure 3.62.

Figure 3.62 *Participants' Writing Weaknesses*



Question 3. The majority of participants said that they had acquired some strong points in their writing. Table 3.36 displays these advantages.

Table 3.36 *Participants' Writing Strengths*

Category	Sub-category	%
Writing structured paragraph		15
Knowing how to evaluate my writing		05
Gaining clarity in my writing		05
Improved in other modules		05
Avoiding mistakes		13
Feeling confident		03
Making clear difference between descriptive and narrative genre		02
Vocabulary improvement	Learning new words	12
	Appropriate use of words	05
Mechanics improvement	Punctuation	02
	Capitalization	05
Language use improvement	Sentence constructions	05

	Tenses	02
	Grammar use	08
Ideas' improvement	Ideas' quality	08
	Ideas' combination	05

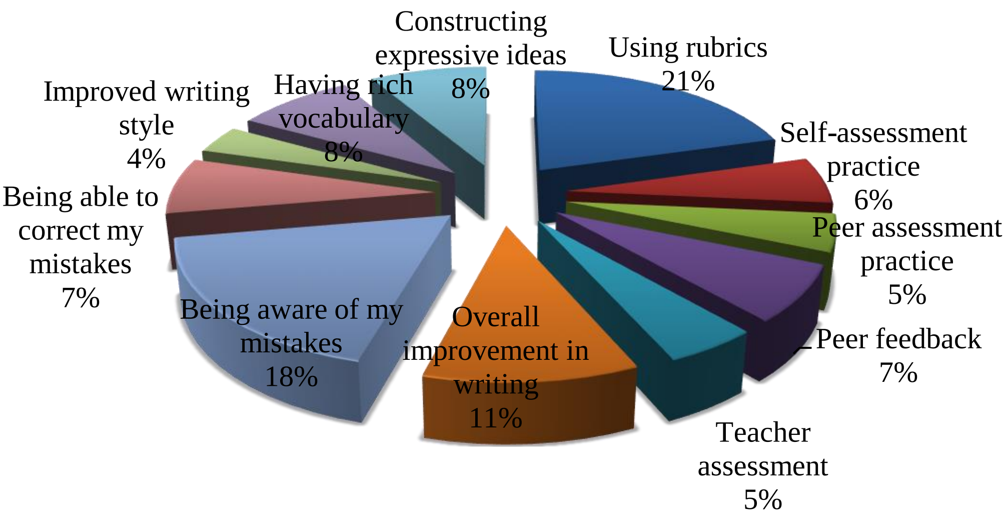
Question 4. All participants acquired self-confidence to a certain extent (64%) or a just a little (36%).

Question 4.1. This self-confidence varied from one genre to another. It is mostly (68%) for both genres followed by the narrative one (23%), and the descriptive one (9%).

Question 4.1. This self-confidence varies according to genre. It is primarily for both genres (68%), followed by narrative (23%), and then descriptive (9%).

Question 4.2. This self-confidence had been acquired from different sources presented in Figure 3.63.

Figure 3.63 *Sources of Gaining Self-confidence in Writing*



Question 5. More than half of the participants (64%) thought that they still need a little teacher guidance to write, and others (12%) thought that they need much guidance while the remaining percentage (24%) thought that they did not need that guidance.

Question 6. More than half of the participants (88%) could compose other types of paragraphs. Here are examples of their answers:

1. ‘I have the basics, so I can write any type just applying its characteristics because they share almost the same points with the descriptive and narrative genres.’
2. ‘I can write any type of paragraph using the rubric because it contains the basics of any piece of writing.’
3. ‘May be I would do it because the structure of the paragraph is the same and I already know how to organize ideas and combine them’
4. ‘Now, I solved many problems in my writing, so I can write easily those types’
5. ‘If I do some practice using the rubric, I will definitely write them in a perfect way.’

Question 7. All participants stated that they are now revising what they have written.

Question 8. All participants, except for four, indicated they improved their paragraph length in comparison to ones written at the beginning.

Question 9. To examine their awareness toward their writing, the participants were asked if they were able to compare their writings before and after taking part in writing classroom assessment. The results obtained are presented in Table 3.37.

Table 3.37 *Participants’ Writing between the Beginning and the End of the Intervention*

At the beginning			At the End		
Category	**Sub-category**	**%**	**Category**	**Sub-category**	**%**
Short paragraph		03	Paragraph developed		18
Unclear writing		15	Clear writing		17
Unstructured writing		07	Structured writing		08
A lot of mistakes		13	fewer mistakes		20
Repetition of words		05	Rich vocabulary		05
Irrelevance to the topic		02	/		/
Problems with language use		09	Language use improved		05
Problems with	**Spelling mistakes**	05	Mechanics	**Capitalization**	05

mechanics	**Capitalization**	05	improved	**Punctuation**	03
	Punctuation	13			
Sentence problems	**Poor constructions**	05	Sentence constructions improved		07
	Different problems	05			
Problems in ideas	**Not expressive**	06	Ideas improved	**Expressive**	07
	Not combined	07		**Connection between ideas**	05

3.3.3 Writing Working Portfolio Analysis

The primary goal of this project is to provide participants with an opportunity to review their written works in order to examine their reactions to their writings, self- and PA. Furthermore, this may help in boosting their awareness of their writing and writing assessment progress. The portfolio project was divided into two sections. The first was divided into four sections: participant writing, self-assessment, APA, and rewriting. It is important to remember that the participants only rewrote when they chose to. Because not all participants had improved their writings, the section on rewriting was left out of the probability analysis. The second section of the portfolio focused on the participants' opinions on the project in connection to their writing.

3.3.3.1 Analysis of the First Part of the Writing Working Portfolio

a.Participants' Works Ordered. The outcomes of the orders placed by participants for their works are displayed in Table 3.38. To grasp the content of this table, three key concerns must be addressed right away. First, the probabilities of combining three pieces of writing, sorting them from best to worst, are found to be six, highlighting the fact that the number of practices was three times for each genre. The letters (D) and (N) stand for descriptive and narrative genres, respectively. The sum of all six possibilities for each level, as well as the sum of all levels, must equal 38 participants.

1. Probability 1 (P1): D1/D2/D3
2. Probability 2 (P2): D1/D3/D2
3. Probability 3 (P3): D2/D1/D3
4. Probability 4 (P4): D2/D3/D1

5. Probability 5 (P5): D3/D1/D2
6. Probability 6 (P6): D3/D2/D1

A, B, and C are attributed to pieces ranging from the best to the poorest one.

1. A is the best written piece/performance according to the participant.
2. B is the average written piece/performance according to the participant.
3. C is the poorest written piece/performance according to the participant.

Table 3.38 *Participants' Ordered Works*

Practice	Genre	Probabilities		Participants' number		
				Level		
				A	**B**	**C**
Writing	**Descriptive**	P1	D1/D2/D3	P1+P2=7(D1)	P3+P5=8(D1)	P4+P6=23(D1)
		P2	D1/D3/D2			
		P3	D2/D1/D3	P3+p4=12(D2)	P1+P6=20(D2)	P2+P5=6(D2)
		P4	D2/D3/D1			
		P5	D3/D1/D2	P5+P6=19(D3)	P3+P4=10(D3)	P1+P3=9(D3)
		P6	D3/D2/D1			
	Narrative	P1	N1/N2/N3	P1+P2= 10(N1)	P3+P5=11(N1)	P4+P6=17(N1)
		P2	N1/N3/N2			
		P3	N2/N1/N3	P3+P4=16(N2)	P1+P6=12(N2)	P2+P5=10(N2)
		P4	N2/N3/N1			
		P5	N3/N1/N2	P5+P6=12(N3)	P3+P4=15(N3)	P1+P3=11(N3)
		P6	N3/N2/N1			
Self- assessment	**Descriptive**	P1	D1/D2/D3	12(D1)	8(D1)	18(D1)
		P2	D1/D3/D2			
		P3	D2/D1/D3	10(D2)	18(D2)	10(D2)
		P4	D2/D3/D1			
		P5	D3/D1/D2	16(D3)	12(D3)	10(D3)
		P6	D3/D2/D1			
	Narr	P1	N1/N2/N3	10(N1)	8(N1)	20(N1)
		P2	N1/N3/N2			

		P3	N2/N1/N3	12(N2)	17(N2)	9(N2)
		P4	N2/N3/N1			
		P5	N3/N1/N2	16(N3)	13(N3)	9(N3)
		P6	N3/N2/N1			
Anonymous Peer Assessment	**Descriptive**	P1	D1/D2/D3	9(D1)	9(D1)	20(D1)
		P2	D1/D3/D2			
		P3	D2/D1/D3	19(D2)	13(D2)	6(D2)
		P4	D2/D3/D1			
		P5	D3/D1/D2	10(D3)	16(D3)	12(D3)
		P6	D3/D2/D1			
	Narrative	P1	N1/N2/N3	9(N1)	12(N1)	17(N1)
		P2	N1/N3/N2			
		P3	N2/N1/N3	16(N2)	12(N2)	10(N2)
		P4	N2/N3/N1			
		P5	N3/N1/N2	13(N3)	14(N3)	11(N3)
		P6	N3/N2/N1			

Table 3.38 shows that in writing practice, the number of good written pieces increases while the number of poor written pieces decreases, and the same is true for student assessments, self- and PA. The written pieces migrating from level C tend to go to level B and are less likely to go to level A. The narrative genre experiences the same pattern, but to a lesser extent. In terms of self-assessment practice, participants strove to enhance their practices over time and reached soon at practicing that accurately, as opposed to APA, where the quality was less significantly improved.

b. Participants' Justifications of the Works Ordered. After ranking their works from best to worst, the participants were asked to defend their ranking. Tables 3.39, 3.40, and 3.41 show their responses. Percentages are chosen rather than raw numbers since many participants left several comments, implying that percentages are more significant than raw numbers in this circumstance. Percentages were generated by using Excel software and computing the percentages of the categories combined, followed by the percentages of the sub- and sub-sub-categories combined. This process did not include PA of the descriptive genre, where even the sub-sub-category is included with the categories, because they looked to be distinct categories rather than sub-sub-categories.

Writing Practice. The categories developed for both levels, poorest and best, might be classified as dichotomies such as written carelessly/ written carefully. Aside from this dichotomization, the categories that cannot be developed as such included ‘ignorance of the writing criteria’ for the lowest level and ‘learned from previous writings’ and ‘learned from the descriptive genre for the best level. The sole category in common between the two levels, poorest and best, is ‘according to peer assessment’, which suggests that the participants picked the best and worst work based on PA.

Table 3.39 *Writing Practice of the Descriptive and Narrative Genre*

Practice	**Level**	**Category**	**Sub-category**	**Sub-sub-category**	**%Descriptive**	**%Narrative**
Writing Practice	Poorest	Ignorance of the writing criteria			02	/
		Disrespecting the genre’s features			11	15
		Lacking practice			03	02
		Problems with the topic			01	02
		Carelessness			08	04
		Following peer feedback assessment			01	/
		Inability to narrate			/	08
		Disrespecting the writing criteria			73	69
		Disrespecting the writing criteria (Theme)	Without details		21	08
			With details	Problems in paragraph form	13	14
				Problems in paragraph title	04	03
				Problems in paragraph length	/	08
				Problems in mechanics	10	05
				Problems in Vocabulary	06	06
				Problems in language use	19	11
				Problems in organization	08	14
				Problems in content	19	31

	Best	Respecting the genre's features			18	20
		Learning from previous writings			07	07
		Learning from descriptive genre			/	02
		Less mistakes			05	04
		Written carefully			07	/
		Topic preference			10	14
		Following peer feedback assessment			02	/
		Having ability to narrate			/	04
		Respecting the writing criteria			51	49
		Respecting the writing criteria (Theme)	Without details		35	15
			With details	Convenient Paragraph form	14	/
				Suitable Paragraph title	04	/
				Correct mechanics	03	/
				Rich vocabulary	10	11
				Appropriate language use	/	15
				Appropriate content	24	41
				Appropriate organization	10	18

Table 3.39 shows how participants justified their choices for the best and poorest written. This seems to be based mostly on writing criteria; disrespecting the writing criteria for the poorest pieces and respecting them for the best ones. This category is followed by dis/respecting the genre's features and written carelessly/ written carefully. The remaining categories appear to be less important. For best writing practice for the narrative genre, specifically 'respecting the writing criteria' category, participants did not focus on technical aspects of writing as they did in the descriptive genre, but instead on significant characteristics such as structure and substance, or content.

Self-assessment Performance. The participants' poorest self-assessment performance was primarily justified by 'ignorance of the rubric criteria', followed by 'overestimation of writing', 'hardness in performing self-assessment' (for descriptive genre only), 'the negative impact of poor writing on self-assessment' (for the descriptive genre only), 'carelessness' (for narrative genre only) and 'underestimating my writing while self-assessment' (for narrative genre only). Participants justified the best self-assessment performance primarily 'appropriate use of the rubric' followed by 'self-assessment done objectively and fairly', 'best work led to objective self-assessment' and 'poor writing led to objective self-assessment'. Further details are presented in Table 3.40.

Table 3.40 *Self-assessment Performance of the Descriptive and Narrative Genres*

Practice	**Level**	**Category**	**Example (The examples of the common categories are extracted from the answers of the descriptive genre section)**	**%Descriptive**	**% Narrative**
Self-assessment Performance	Poorest	Mishandling the rubrics	'I disrespected the descriptive levels' 'I didn't know how to use the rubric properly'	29	48
		Overestimating my writing when self-assessing	'I put excellent for use of vocabulary but in reality it is average' 'I assessed organization excellent instead of poor in reality' 'I didn't assess myself in a good way and in some points I was subjective'	33	35
		Hard performance of self-assessment	'I was new with self-assessment' 'It was not easy to self-assess my writing at the beginning'	22	/
		Negative impact of self-assessment on writing- Producing poor writing	'I didn't pay attention while assessing my poor paragraph' 'I assessed my worst paragraph in a harsh way'	14	/
		Carelessness	'I was tired while assessing my writing, so I didn't focus on the assessment criteria' 'I assessed the paragraph in the same way I write; It was done rapidly'	/	14

		Underestimating my writing when self-assessing	‘ I underestimated my writing a lot’	/	03
		No answers	/	02	/
	Best	Appropriate use of rubrics	‘Because I understand the descriptive levels and how to use the rubric’ ‘Because I applied the rubric correctly’	60	26
		Objective and fair Self-assessment	‘I was fair in assessing my writing’ ‘I did it objectively looking for all the details in the rubric’	29	58
		Best work boosting to objective self-assessment	‘I was satisfied with my writing, so I assessed it carefully’	08	16
		Poor writing boosting to objective self-assessment	‘It was not good, that’s why I decided to assess it objectively’	03	/

Peer Assessment Performance. The poorest peer assessment is seen as so because of the incorrect peer assessment, uninteresting peer feedback, and overestimation and underestimation of the peer’s writing. On the other hand, the best peer assessment is mainly supported by objectivity and fairness of delivering peer assessment, helpful peer feedback, respecting rubrics, and providing correct peer assessment. Further details are displayed in Table 3.41 below.

Table 3.41 *Anonymous Peer Assessment Performance of the Descriptive and Narrative Genres*

Practice	**Level**	**Category**	**Sub-category**	**Sub-sub-category**	**Example(The examples of the common categories are extracted from the answers of the narrative genre section)**	**%Descriptive**	**% Narrative**
Peer	Poorest	Uninteresting peer feedback			‘he/she didn’t write a lot about my paragraph. He/she only said that I had bad writing’ ‘My classmate didn’t criticize me	19	21

				enough as his remarks was simple not expressive'		
	Overestimating my writing			'The assessment was good but I know that my writing was poor, so I think my classmate overestimated me' 's/he went easy on me. I think s/he overestimated my writing'	16	11
	Underestimating my writing			'The person that assessed my paragraph underestimated me'	15	11
	Unserious peer assessment			'I think she was not serious when assessing my writing'	/	03
	Problems with the way of giving peer feedback			'the assessor was a little bite rude. He didn't give me advice in a good way'	/	04
	Disagreement between peer and self- assessment			'This is the poorest peer assessment because the assessor didn't give me the right'	/	07
	Incorrect peer assessment			/	/	43
	Incorrect peer assessment (Theme)	Without details		'His assessment was incorrect because my writing was terrific' 'Despite having a lot of mistakes, my classmate said that my writing was good'	/	25
		With details	Mishandling rubrics	' It was a bad assessment because he didn't follow the rubric' 'It was not that fair and the assessor didn't respect the rubric'	18	33
			Lack of objectivity	'My classmate's assessment was brief and lack of objectivity'	32	42
Best	Fair and objective peer assessment			'My classmate's judgement was correct and fair' 'I got fair assessment and I think that the assessor gave me what I deserve in a fair and objective way'	34	45

	Helpful peer feedback	‘The remarks my classmate provided were very constructive and helpful’	18	13
	Respecting rubrics	‘My classmate did respect the rubric in a good way’ ‘I was please because the assessor respected the norms and assessed my writing in a perfect way’	15	10
	Correct peer assessment	‘He pointed out all my mistakes’ ‘He said the truth about my mistakes in my paragraph’	12	04
	Encouraging peer feedback	‘He gave me solutions how to improve my writing’	09	07
	Impact of providing peer feedback	‘I liked the way of providing remarks because they encouraged me a lot’	06	/
	Agreement between self- and peer assessment	‘I and my classmate did nearly the same assessment’ ‘My classmate and I agreed upon the assessment’	03	15
	Selecting best assessment following rubric use	‘The majority of the assessments in the rubric were good and excellent’	03	06

Rewriting Practice. The participants justified the best and the poorest of their rewritings in many ways. The poorest rewritings were classified as so because the participants did not do any major changes on the writing while for the best ones were labeled as such because the participants tried to correct the mistakes they made.

Table 3.42 *Rewriting Descriptive and Narrative Genre Writings*

Practice	**Level**	**Category**	**Sub-category**	**%Descriptive %**	**%Narrative**
Rewr	Poore	No major changes		19	20
		Carelessness		16	05

Best	Having problems with the topic			03	08
	Disrespecting the rubrics’ criteria			06	/
	Excluding peer feedback			03	/
	Unsatisfied with my rewriting			03	13
	Disrespecting the genre’s features			/	13
	Following peer assessment			/	03
	Lacking specific details			50	38
	Lacking specific details (Theme)	Problems with paragraph length		06	06
		Limited vocabulary		25	06
		Problems in punctuation		06	07
		Problems in different types of sentences		12	27
		Problems with connectors		13	27
		Weak connection between ideas		13	13
		Weak paragraph structure		06	/
		Problems with supporting details		19	07
		Problems with organization		/	07
	Satisfied of my rewriting			17	09
	Respecting the genre’s features			06	09
	Spotting and correcting mistakes			11	16
	Rubric help in rewriting			06	04
	Following peer assessment			02	/
	Helpful peer feedback			/	04
	Liking writing topic			04	/
	Improved writing with practice			/	02
	Significant improvement			54	56
		Without details		32	/
		With details	Rich vocabulary	14	04
			Grammar improved	27	08
			Punctuation improved	/	12
			Using different types of sentences	04	14

	Significant improvement (Theme)		Respecting paragraph structure	09	18
			Respecting paragraph form	/	10
			Improved use of connectors	04	04
			More details	05	10
			Improved coherence	/	04
			Improved topic sentence	/	04
			Improved concluding sentence	05	04
			Relevance to the topic	/	06
			Ideas improved	05	10
			Handwriting improved	05	04

As presented in Table 3.42, poorest rewritings are mostly justified by lacking specific details for both genres. To elaborate on this category for the descriptive genre, it includes restricted use of vocabulary followed by problems in the supporting details, connectors, and connection between ideas. For the narrative genre and in the same context of lacking specific details, that is basically marked by problems in different types of sentences and connection between ideas. In addition to lacking specific details, ‘no major changes’ comes in the second position for both genres followed by ‘carelessness’ for the descriptive genre and ‘dissatisfaction and disrespecting the genre features’ for the narrative genre.

Best rewritings, on the other hand, are justified by significant improvement in writing both genres. To elaborate on this category for the descriptive genre, it includes ‘improved grammar’ and ‘using rich vocabulary’. For the narrative, in the same context of significant improvement, this is basically marked by ‘respecting the paragraph structure’, ‘using various types of sentences’ and ‘correct use of punctuation marks’. In addition to significant improvement, ‘satisfied of my writing’ figures out in the second position followed by ‘recognizing and correcting my mistakes’ for the descriptive genre whereas for the narrative genre that is mainly marked by ‘recognizing and correcting my mistakes’ followed by ‘satisfied of my writing’ and ‘respecting the genre features’.

3.3.3.2 Analysis of the Second Part of the Writing Working Portfolio. After ranking their works from the best to the poorest one, the participants were asked about their perceptions of experiencing working portfolio. The data collected in this section are displayed in Table 3.43.

Table 3.43 *Participants’ Perception of Writing Working Portfolio*

Theme	Example	%	Category	Example	%
Motivating students	‘I am motivated by seeing all my works at once.’ ‘Seeing the difference between my writings gives me a message that I can progress and develop it. It is possible and I can do it.’	80	Knowing their strengths	‘Knowing that I have acquired some criteria of good writing gives me self-confidence.’	17
			Knowing their weaknesses	‘I spotted all my mistakes. The only thing to do now is to correct them all.’	63
Developing self-reflection and choice	‘This experience definitely changed my view of how to write a paragraph.’ ‘It is good to see all my writing at once because I can compare them to know the good criteria of writing.’				43
Promoting autonomy and ownership	‘I can understand that to improve my writing depends on myself ’ ‘The portfolio teaches me to depend on myself in my writing.’				38

To a large part, the working portfolio is considered as a way that stimulated participants to reconsider their writing as well as generating self-reflection and choice in them to boost their autonomy and ownership.

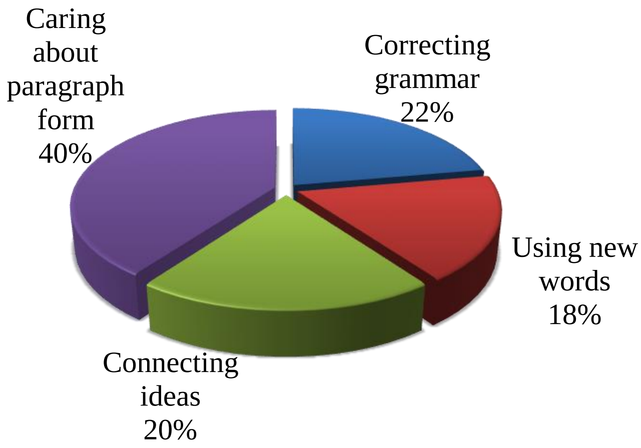
3.3.4 Analysis of the Questionnaire Administered to the Control Group and Analysis of the Post-test

3.3.4.1 Analysis of the Questionnaire Administered to the Control Group

Question 1. Less than half of the students (42%) thought their writing had improved, whereas the remaining students (58%) did not.

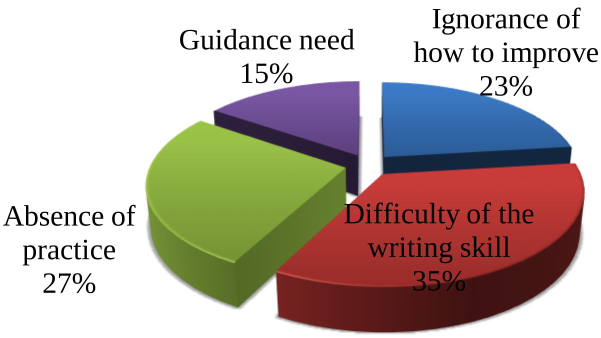
Question 1.1. When asked how they accomplished their development, the participants gave various answers, as seen in Figure 3.64.

Figure 3.64 *Ways of Achieving Progress in Writing*



Question 1.2. 72 participants who answered ‘No’ had explained why they did not improve their writing. Their answers are grouped into four main categories, as shown in Figure 3.65.

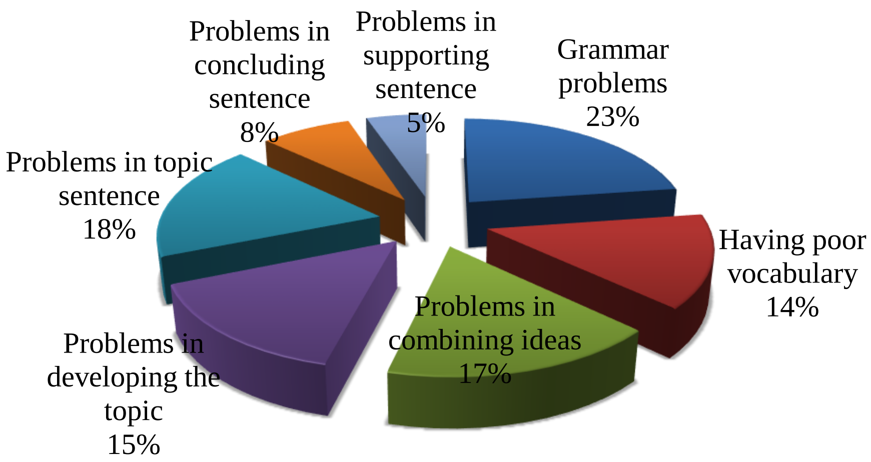
Figure 3.65 *Causes of Non-progress in Writing*



Question 2. When questioned about their motivation to improve their writing; (41%) stated they had ‘a little’ motivation and (13%) responded by ‘Yes’ while the others (46%) responded by ‘No’

Question 3. All participants admitted to still having weaknesses in their writing. These problems are presented in Figure 3.66.

Figure 3.66 *Participants' Writing Weaknesses*



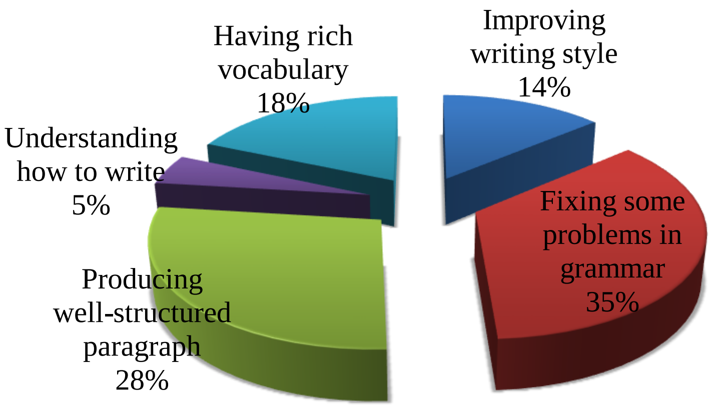
Question 4. More than half of the participants (62%) indicated they improved their writing and corrected some of the recurring errors they made. Table 3.44 shows the students' writing strengths.

Table 3.44 *Participants' Writing Strengths*

Category	%
Structured paragraph	15
Avoiding mistakes	13
Vocabulary improved	09
Grammar improved	28
Ideas improved	35

Question 5. About (32%) believed they had developed 'a little' confidence in their writing while (60 %) thought they had not, and just (8%) confirmed they had grown self-confidence. For those who answered by 'a little' or 'Yes', they explained their answers differently as displayed in Figure 3.67.

Figure 3.67 *Sources of Gaining Self-confidence in Writing*



Question 6. (61%) of the participants admitted to being anxious and stressed while writing, and (26%) experienced 'a little' anxiety and stress. The remaining (13%) did not have anxiety or stress at all.

Question 7. More than half of the participants (32%) believed that they still need a little teacher guidance to write and (52%) thought that they definitely need that guidance while (16%) believed they do not.

Question 8. Approximately 46% of participants still do not review what they write, whereas 54% do.

Question 9. To elicit their awareness of their writing, the participants were asked whether they observed any difference in their writing between the beginning and the end of the year; 39% of the participants recognized the difference while the others not.

Table 3.45 *Participant Writing between the Beginning and End of the intervention*

At the beginning		At the End of intervention	
Category	**%**	**Category**	**%**
Unstructured writing	07	Structured writing	03
A lot of mistakes	25	Less mistakes	20
Poor vocabulary	18	Rich vocabulary	22
Problems with grammar	28	Accurate use of grammar	39
Problems with spelling mistakes	10	Improved spelling mistakes	06

Problems in combining ideas	12	Appropriate combination of ideas	10

3.3.4.2 Analysis of the Post-study Test.

Content analysis of the post-test yielded various results displayed in Table 3.46. For comparative reasons, the percentages obtained before intervention are also mentioned.

Table 3.46 *Analysis of the Written Post-test*

Errors and problems			Presence			
Type	**Category**	**Sub-category**	**Exp. group (Before intervention)**	**Exp. group (After intervention.)**	**Cont. group (Before intervention)**	**Cont. group (After intervention)**
Substance	Spelling	Typographic & cognitive	91%	12%	72%	38%
	Punctuation	/	100%	15%	100%	40%
	Capitalization	/	78%	25%	72%	28%
Textual	Lexical	Lack of vocabulary	82%	28%	78%	58%
		Contextual meaning	48%	08%	43%	23%
		Foreign words	05%	01%	10%	01%
	Grammatical	/	100%	33%	100%	42%
	Syntax	Sentence errors	62%	18%	70%	34%
		Inter-sentence errors (cohesion)	72%	21%	80%	35%
Discourse	Coherence	Topical coherence	74%	21%	80%	32%
		Relational coherence	62%	18%	58%	21%
Paragraph Structure Problems	Topic sentence		100%	18%	100%	48%
	Supporting Sentences		71%	43%	85%	51%
	Concluding Sentence		72%	01%	78%	18%
Form and Appearance Problems	Paragraph Length		82%	03%	48%	12%
	Paragraph entitlement		87%	00%	90%	55%
	Block format		88%	00%	78%	12%
	Indentation		78%	00%	61%	08%
	Handwriting		04%	00%	02%	00%

The post-study test analysis revealed that there are improvements in the writings of both groups but with various degrees. The experiment group had benefited from their experience with AfL and learned how to improve their writing much more than did the control group who did not receive any special treatment.

3.4 Chapter Summary and Conclusion

The analytic procedures implemented in this research revealed that the 77 participants had numerous writing issues, namely related to vocabulary, language use, mechanics, cohesion and coherence. The responses supplied in the pre-study questionnaire and the analysis of the written pre-test are used to cross-check this fact. The problems were not only encountered by the research participants, but also by two other groups of the same population. This is to ensure the validity of the sample of the study and shed light on the writing difficulties encountered by first year English degree students at the department of English in the University of Algiers 2.

An in-depth examination of the various causes behind this, as well as the attempts to adopt assessment processes, indicates that students in high schools were not given the chance to participate in the classroom because of, essentially, program constraints. It was also revealed that, except for the few truncated attempts, the student participants were not involved in undertaking student assessment, self- and PA. Furthermore, many additional factors for students' poor writing were recognized, such as inadequate vocabulary, grammatical challenges, ignorance of the writing standards, and individual efforts toward writing improvements as not being directed and encouraged.

The participants basically demonstrated that they had little experience with rubrics and had never been given the opportunity to be involved or proactive in the classroom. Despite the obstacles encountered at the beginning of the study, they effectively used the rubrics and did self- and APA after becoming involved. Those impediments were justified by a lack of desire, a lack of experience with writing classroom assessments, and their reliance on instructors as they were in classroom, which supports teacher-centeredness. In terms of the purposeful examination of the participants' works, the analysis concluded to varying degrees of progress in their writing as a result of their response and efforts. Furthermore, the participants valued the usage of rubrics since they found it inspiring and encouraging, as well as a means to convey their strengths and weaknesses in their works. However, they underlined the need of using the scoring rubrics but after using the analytic instructional ones to avoid overestimating their works. They stressed using scoring rubrics because they appraised the importance of scores and using that after investing in qualitative feedback to give a space for the feedback to exercise its role and impact. As for the writing portfolio, its importance lies in enabling participants to backtrack their development and remind them of their problems to consider and work on them harder.

The T test analysis using SPSS version 26 revealed that there is a strong significance in terms of writing progress between the two groups, but the experiment group highly outperformed the control group. It is worth being mindful that the control group outperformed the experiment one at the beginning of the intervention. The main findings on writing development for the experiment group were making significant progress in vocabulary use, language use, cohesion and coherence in addition to gaining self-confidence, feeling less anxious while writing. Furthermore, participants understood that revision after writing is a must and more importantly they learned how to do that, focusing on the criteria of rubrics. Regarding writing content improvements, a great majority of participants knew how and what to improve in their writing, through focusing on the writing criteria and comprehending what aspects they needed to consider making that development. Participants showed that making progress in the descriptive genre was easier than in the narrative genre, and they pointed to the need of more sessions to make better development in this genre. In the control group, the participants reported a lack of self-confidence and a high level of stress, particularly while completing their writing tests. Moreover, they have shown more writing problems at both the content and form despite having shown higher writing levels at the start of the intervention.

4

CHAPTER FOUR: DISCUSSION OF THE FINDINGS, RECOMMENDATIONS FOR FURTHER RESEARCH, AND PRACTICAL PEDAGOGICAL IMPLICATIONS

After presenting the collected data, the discussion in this chapter proceeds for each research question in light of the results and the existing literature. This is the first part of this chapter. The second part addresses the major recommendations for further research in which the main factors affecting AfL implementation in relation to the underlying procedures as operationalized in this study are scrutinized. This includes, rubrics, self- and PA, and overall AfL classroom.

4.1 Discussion of the Findings in Relation to the Research Questions

Before embarking on the discussion proper, it is preferable to highlight the major limitations within which this discussion takes place in order to avoid drawing attention away from the primary objective meant to be addressed in light of this work. The first important constraint is theoretical in nature. As we stated in the first section of the literature review chapter, AfL lacks a distinct theory and has yet to be theorized. However, certain efforts were made to promote AfL implementation, particularly the ten principles communicated by the ARG in order to effectively activate it in the classroom. Many theories, including socio-constructivism, cognition, and motivation, are just stated without explanation to support AfL. Furthermore, some experts believe that the findings of prior investigations are a precursor to support AfL. That is proposed in order to provide flexibility for researchers in order to motivate them to engage in the AfL research area, as well as to assist in clearly operationalizing and defining this idea in many ways, in response to the objectives delineated by the study to be conducted. The second constraint to identify is the lack of similar research in the EFL environment in general, and their absence in the Algerian setting in particular. This may impede and disorient researchers since they are focused on replicating already done studies rather than trying to undertake a new study developed by their own efforts. In this spirit, we strongly urge researchers to participate in this study field because there are numerous chances to grab various implementing techniques and processes for concretizing AfL on terrain. Because the notion of AfL is not just for summative purposes, it is more concerned with feasibility and practicability than replication.

4.1.1 Writing Difficulties of First Year EFL Degree Students

To answer the first research question, which served as an exploratory stage prior to the intervention proper, the major results used for the discussion are displayed in Table 4.1, indicating which data gathering tool and precisely which section (s) and item (s) were used. To be clear, the discussion will concern first year students' EFL writing problems, spotting them and explaining the major causes which generated them.

Table 4.1 *Results Used to Discuss the First Research Question*

Construct	Data gathering tool	Data used for the discussion
EFL writing problems	Pre-study questionnaire	Section 2: Q1 (Q1.1 & Q1.2), Q2, Q3 Section 3: Q1, Q2, Q3, Q4, Q5, Q6 Section 4: Q1, Q2, Q3, Q4, Q5 (Q5.1 & Q5.2), Q6 (Q6.1 & Q6.2), Q7 (Q7.1 & Q7.2) Section 5: Q1, Q2, Q3, Q4, Q5
	Written pre-test	Content analysis of the written pre-test
	Follow-up section of the rubric in self-assessment	Answers to whether 'you are satisfied with your writing or not and why'

The student participants of this study revealed that English writing is difficult. This included both experimental and control groups, as well as two other groups drawn at random from the same population. When identified, the difficulties revealed to be grouped into five categories: mechanics, vocabulary, language use, content, and organization.

The most common issues in mechanics were with punctuation marks, capitalization, and spelling errors. As for vocabulary, the difficulty lies in lack of vocabulary, word-context meaning, and the usage of foreign words. The extracted challenges in language usage covered sentence structures, s-v agreement, tenses, use of articles, pronouns, prepositions, and negatives, as well as difficulties in meaning at the sentence level. For organizational issues, these included difficulties in expressing ideas, making them explicit, and correctly integrating them. The descriptive genre was essentially marked with issues of space order and/or order of significance, whereas the narrative genre was specifically characterized with the issue of time order. The major problems with content included paragraph structure, within which topic sentence, supporting sentence, and concluding sentence were spotted, and relevance to the topic.

The findings are consistent with previous research on EFL writing challenges undertaken in the Algerian setting, such as Ourghi (2002), Hamzaoui-Elachachi (2006), Hamzaoui-Elachachi (2010), and Chelli (2013). To report on the difficulties detected in first-year students' writings, Ourghi (2002) found out that the majority of first-year students (80%) had low-intermediate level and only few had high-intermediate level (20%). The research findings on low-intermediate level writing difficulties revealed deficits in the use of syntactic structures, writing mechanics, vocabulary, composing effective writing strategies, using grammar, and cohesion and coherence (Hamzaoui-Elachachi, 2006; Hamzaoui-Elachachi, 2010).

The students in this research were conscious of their low level of written expression and showed readiness to change. They explained their incapacity to write correct English, their anxiety and stress, and their lack of self-confidence when writing, as well as their failure to revise their work. This was mostly due to a lack of vocabulary, grammar issues, inability to communicate and think in English, and writing composing strategies because the writing skill was regarded to be challenging and complicated. These difficulties appear to be also widespread throughout the Arab world, rather than being exclusive to the Algerian context. Many academics in other EFL contexts have written on this topic, identifying comparable challenges; mentioning some of them like Tunisia (e.g., Enneifer, 2021), Morocco (e.g., Rakrak, 2020), Egypt (e.g., Ahmed, 2010), Saudi Arabia (e.g., Mohammed & Hazarika, 2016), and Jordan (Ibnian, 2017). Beyond the Arab world, English writing issues are also found in a variety of non-Arab contexts, including China (e.g. Zou, Kong, Lee, 2021), Spain (e.g. Montaner-Villalba, 2019), and Turkey (e.g. Kirmizi & Aydin, 2019).

Rather than focusing on the complexity of the writing skill per se, we found that writing challenges were also caused by the teaching paradigm used in the classes in high schools, such as the lack of engaging and reactivating students. Students not only did not receive any feedback from their teachers or classmates, but they were also not given opportunity to reflect on their own work, with the exception of a few rudimentary initiatives. In this context, Chelli (2013) reported that students' compositions were solely graded in exams, with no opportunity for reflection before or during the writing growth process. Furthermore, lack of frequent evaluation was identified as one of the factors for the emergence of writing issues among EFL students, which was justified by large classes and the use of traditional techniques of teaching and assessing writing (Moussaoui, 2012).

At individual preference level, the student participants in this study demonstrated that they did not enjoy writing, did not write in their spare time, and did not revise what they wrote. They were generally uninterested in developing their writing for reason related to motivation and learning practices imposed in classrooms, like working individually rather than collaboratively. Despite this, students were aware of the value of feedback in improving their writing and identifying errors, and they perceived the feedback provided by their teachers in high schools to be insufficiently specific and inadequate in certain circumstances. For further details about feedback provided by high school teachers, see figure 3.4.

We invited participants to envision themselves as instructors and offer techniques to aid their students to write effectively. The outcomes were so interesting. They responded in various ways, each providing effective solutions and advice (For detailed explanation see Figure 3.42). They insisted on assisting students to acquire new vocabulary, immersing them in frequent practice both in the classroom and at home, providing counsel and guidance, and feedbacking regularly their students' works. They also proposed, despite their small number, exposing samples of effective writing for their students, establishing an interactive environment in the classroom, grading their work without putting it in the final grade, allowing students to make mistakes, and writing cooperatively to learn from one another. These were highly practical ideas that were mostly and almost entirely taken into consideration since AfL is essentially stressing all of that. This supports that students are an effective source of feedback when they participate in the discussion (Topping, 2019; Andrade & Boulay, 2003; Bangert-Downs, Kulik, Kulik, & Morgan, 1991; Kulhavy and Stock, 1989; Meyer, 1986).

To maintain the discussion within the study's restrictions and delimitations, we do not claim any generalizability based on the findings, although they may be efficient within the limit measurements of four groups, over 20 groups of the population. The four groups all submitted written pieces for further content examination. However, based on the many studies in the literature on the topic of EFL writing problems, none of which deny the existence of writing problems in the EFL context, we could conclude that first-year EFL students in the population under study had writing problems that required additional interventions to help them improve their writing and overcome their difficulties.

4.1.2 The Impact of Using Writing Rubrics on First Year EFL Degree Students' Writings

To answer the second research question, which intends to investigate the use of rubrics in EFL writing classroom evaluation, the primary results utilized for the discussion are given in Table 4.2, indicating which data collection instrument was used and precisely which part (s) and item (s) were exploited. The discussion focuses on how students use rubrics, how that happens, and if the whole process is beneficial and effective, pointing out any obstacles and making additional recommendations.

Table 4.2 *Results Used to Discuss the Second Research Question*

Construct	Data gathering tool	Data used for the discussion
Writing rubrics' use	Writing rubrics	Qualitative and quantitative/scoring rubrics
	Follow-up sections of rubrics in self- and peer	Answers to whether the student was satisfied with using the rubric or not and why (Q1: Q1.1 & Q1.2)

	assessment	and whether the rubric helped the students to assess their classmates' writings and why (Q1: Q1.2)
	Researcher journal report	Section 4 (4.1 & 4.2)
	Post-study questionnaire (Part I)	Section 1: Q1, Q2, Q3, Q4, Q5, Q6, Q7, Q8, Q9, Q10, Q11 Section 2: Q3 Section 3: Q3

In this study, the use of rubrics to improve students' writing was an exceptional experience because they were unfamiliar with such tools and classroom assessment in general. Despite having been piloted and familiarized to participants, the students did not respond fruitfully to utilizing rubrics at the start of the intervention because they were processing the issue, indicating a latent phase. Following this period, they began providing comments and disclosing further questions. With experience, their questions faded and they successfully comprehended the applications and goals of rubrics. The significance of rubrics for students may be summed up in four points: (1) Rubrics aid in spotting errors, assessing writing, improving writing, and communicating students' strengths and weaknesses to teachers; (2) rubrics facilitate assessment; (3) rubrics define rules for writing and assessing effectively, and even define writing quality; and (4) rubrics teach the rules of writing and assessment.

As for scoring rubrics, students concluded that while marks were necessary, they should not be accompanied with rubrics right from the beginning since this might lead to students overestimating their work and lose interest in developing their writing and resolving their difficulties. That was cross-checked with students' responses in the pre-study questionnaire when questioned about the value of feedback and marks in improving their performance, and they said that both were important.

In this part, we will address the significant consequences of students adopting writing rubrics.

a. Using writing rubrics motivates students, increases self-confidence, decreases anxiety, and boosts self-efficacy. The student participants perceived their experience with rubrics as an intriguing project because they were inspired to improve their writing, felt the need to rewrite their work to recognize and solve their problems, felt less worried when writing, and felt more confident. They each stated it differently. One student wrote, ‘I feel that I am confident in my writing because rubrics helped me a lot. They guided me to see my mistakes and correct them.’ Another one stated that rubrics were tools that should be kept in mind not only temporarily but all the time because they cover all of the fundamentals of writing.

The results obtained go in agreement with some studies like Andrade and Du (2005), Reynolds-Keefer (2010), and Panadero and Romero (2014) who found that using rubrics helped students decrease their anxiety and feel more confident. Furthermore, Andrade and Du (2005) found that employing rubrics for both students and teachers defined the expectations obtained from accomplishing activities and grading the assessment to activate the feedforwarding mechanism; explaining how to improve performance and enabling students to reflect on their work (Truemper, 2004). Panadero, Jonsson, and Strijbos (2016), on the other hand, see that rubrics emphasize primarily the method of assessing either the result, the process, or both, rather than offering a set of instructions on how to do the work. We found in this study that rubrics catalyzed and stimulated students’ cognitive abilities toward writing, as a recursive process, by supporting their composing strategies to feedback and feedforward their writing progress.

In addition to being an important tool in mediating student performance improvement, rubrics may play a key role in promoting student self-efficacy, lowering fear, increasing transparency, and assisting in the feedback process (Panadero & Jonsson, 2013). As a result, rubrics are used not only to express the requirements, but also to elevate students’ writing aspirations to higher levels. On the other hand, Covill (2012) believes that rubrics, when used in accordance with social cognitive theory, can promote self-efficacy and hence motivation and writing performance. It may also help in constructing students’ conventions about the target goals for their writing and it may scaffold for constructing meaning. For example, rubrics used in the study assisted students in distinguishing between descriptive and narrative genres, in addition to evaluating the basics of writing. In this regard, students acknowledged that creating other genres, such as opinion or process paragraphs, needs learning the foundations of writing as well as respecting the genre, rather than viewing each genre as distinct from others in the writing field.

b. Writing rubrics reflect students' strengths and weaknesses. Students saw the use of rubrics as a tool to reflect their strengths and flaws. During their writing assessment engagement, we have noticed that students had tried to expose their inquiries and interact with the teacher and their peers, exposing their worries about writing and rubrics as assessment tools that enhance their progress. They realized that the purpose of the rubric was not only to apply the criteria correctly, but also to guide them to make realistic progress by providing a venue for debate. In agreement with this, Crawford (2001) sees that rubrics allow both students and teachers to evaluate the expectations taken from every assignment in order to have a clear knowledge of the end output. Yet, neglecting the writing criteria was primarily found to be the most significant barrier to students' failure to revise their writings (White, 1994; Cox et al., 2015).

Students may distinguish their flaws from their strengths by employing rubrics, and vice versa. Instead of focusing just on one area and treating it as their whole problem, they understood that their issues were far broader. To put it another way, at first, students concentrated mostly on grammar, attempting to identify all of their problems within this area, but after using rubrics, they realized that writing involves more than just generating error-free writings; there were other factors to consider. In terms of their strengths, students had improved their language use, which inspired them to learn new vocabulary, create appropriate paragraph structure, be topic-related, and comprehend genre. They also recognized that they needed to work harder to improve a variety of characteristics of language use, such as sentence building and punctuation marks, in addition to improving paragraph structure and trying to learn additional vocabulary (for more details see Figure 3.66 and Table 3.44)

c. Writing rubrics are tools to communicate students' needs to teachers. One of the important effects reported from students' experience with using writing rubrics was considering rubrics as a tool to communicate their needs to teachers. This had been expressed differently like ‘I see that this rubric is a good way to make my teacher know about my problems in writing’ and ‘I found the rubric interesting for me because it gives an idea for my teacher about my writing problems’. However, Covill (2012) claims that most students indicated that being asked to examine their work had impacted their writing skills in positive ways. That may be accomplished by fostering communication between students and teachers through the use of rubrics, which would allow both parties to reflect and react appropriately on the writing stages.

In this context, we found that when students reported their concerns, they informed the teacher, who then looked for further solutions and positive criticism through constructive and elaborated feedback. This is substantiated by the questions students raised in the follow-up part of rubrics, whether in self- or APA. Each time, students reported their writing challenges and identified the areas in which they had improved, as well as their progress along the descriptors’ continuum. They learnt how to improve their writing by accepting responsibility and acknowledging their faults before moving on to correction.

d. Writing rubrics are necessary to student assessment performances; self- and peer assessment. Students could not envisage student assessment performances without rubrics since they were perceived as challenging due to a lack of knowledge about the standards to follow when assessing. Furthermore, they believed that rubrics provided legitimacy and objectivity for evaluation. In support of this, Wang (2016) reported that employing rubrics in self-assessment was more helpful in fostering students’ self-regulated learning. This impact will be extensively discussed when the third and fourth research questions are discussed (Sections 4.1.3 & 4.1.4).

e. Scoring writing rubric supplements positively the use of the instructional qualitative rubrics. Students emphasized the value of scoring rubrics after using the qualitative instructional rubrics for a period of time. This was explained by breaking students’ attachment to scores and allowing them to focus on their growth. In an important AfL project conducted by Black, Harrison, Lee, Marshall, and Wiliam (2004), it was revealed that students given marks saw it as a way to compare themselves to others, whereas students given only comments saw it as a way to help them improve, and the latter group outperformed the former. In this study, students found that scores were also required to validate progress, and students explicitly stated this in the pre-study questionnaire. In accordance with this, Cox et al. (2015) believe that rubrics must be robust enough to assure both marking and offering constructive feedback in order to reflect learning outcomes. In this study, scoring rubrics were shown to motivate students in two separate ways: advancing in reaction to stress and progressing without being stressed. Similarly, Stiggins (2002) highlighted how employing assessment to encourage students for learning, arguing that one approach of maximizing learning was to enhance fear via using assessment because assessment on its own has been viewed as the great intimidator.

4.1.3 The Impact of Performing Self-assessment on First Year EFL Degree Students' Writings

To answer the third research question, which intends to analyze self-assessment performance in EFL writing classroom assessment, Table 4.3 displays the primary results employed for the discussion, stating which data collection method was used and precisely which section(s) and item(s) were examined. The discussion focuses on students undertaking self-assessment, how that was accomplished, and if the entire process was effective and beneficial, highlighting any challenges for future recommendations.

Table 4.3 *Results Used to Discuss the Third Research Question*

Construct	Data gathering tool	Analyzed data used for the discussion
Student self-assessment	Pre-study questionnaire	Section 2: Q2 Section 4: Q6 (Q6.1 & Q6.2)
	Writing rubrics	Analyzing self-assessment done through using rubrics
	Researcher's journal report	Section 2
	Post-study questionnaire (Part II)	Section 2: Q1, Q2, Q3, Q4, Q5, Q6

The primary outcome of completing self-assessment in writing classroom assessment was to encourage students to engage in self-regulated learning. Self-assessment stimulated students to participate in writing improvement by being conscious of taking responsibility of their own writing development and doing so exclusively for themselves, not for their teacher. To emphasize the significance of self-assessment, Panadero and Alonso-Tapia (2013) see that self-assessment is a key component of self-regulated learning, which is defined as "self-generated thoughts, feelings, and actions that are planned and cyclically adapted to the attainment of personal goals" (Zimmerman, 2000, p. 14). On the other hand, Topping (2019) sees that in the longer term, self-assessment "might impact self-management or self-regulation of learning – facilitating continuous adaptation, modification and tuning of learning by the learner, rather than waiting for others to intervene, and this is "much the same as peer assessment" (p. 14). As a result of being self-regulated, students can progress forward and backward while developing skills and knowledge (Bransford et al., 2000).

In fact, student participants initially misunderstood the concept of self-assessment since they were unfamiliar with it. Because they had no classroom assessment experience, they initially believed that self-assessment was only a task to be completed without regard for the significance of the issue. With repeated practice, students realized that this was more than simply a sporadic task, but that it was necessary to participate in further writing progress. When feedbacked their works via self-assessment, they displayed various opinions regarding their works. Some had written briefly about their efforts, appraising them, especially at first, but as time passed, they altered their minds and began detailing their problems as well as spotting their growth and describing how that occurred. They mostly valued the use of rubrics as well as showcasing their queries related to the progress they had made.

In contrast to the notion that self-assessment is just asking students to grade themselves, or self-grading, the students in this study urged not to prioritize self-grading over self-feedback because this might alter the whole process. They did, however, regard self-grading to be a vital step to include in the total performance. Andrade and Boulay (2003) demonstrated in this context that self-assessment could do more by supporting skills and learning development through students' careful reflection on their own work, considering themselves as a source of feedback under appropriate conditions and supports, particularly when it was not interpreted as self-grading. In their study, they see that the relevant support was the instructional rubrics that described good and poor writing, and the appropriate conditions were two self-assessment lessons they devised to help students use a rubric to assess their draft essays. Moreover, Gipps (1994) sees that self-assessment is not pure self-grading more than being linked with one’s self-monitoring to enhance motivation if it is used carefully and conveniently.

Another important result to shed light on is that despite having no ability to self-assess, low-achievers, in particular, had succeeded in doing so to a great extent. This reveals that students could self-assess even with no or just rudimentary skills as suggested by Panadero et al. (2017). These authors emphasized in their meta-analysis that self-assessment, while it can occur in the most elementary form without any help, can also be more broad-ranging and supportive via using assessment tools such as rubrics, training, and/or feedback on self-assessment performance. Despite the difficulty of the exercise, students demonstrated that they can self-assess their work when guided and scaffolded. In comparison to the second student performance, which is PA, we found that students had more difficulties with PA since they felt more responsibility for any score or feedback they provided. This is thoroughly explained in the discussion section related to fourth research question (see Section 4.1.4).

It was also found that some students with poor writing skills tended to exaggerate their work at first, whereas they tried to examine their work objectively after realizing their writing problems. They recognized that self-assessment was for their own benefit and not to demonstrate to the teacher that they were flawless despite the fact that they were not. On the other hand, a lot of good student writers, particularly boys, tended to overestimate their ability and maintained this throughout the whole intervention. This contradicts the findings of Falchikov and Boud (1989) and Prohaska and Maraj (1995), who found that beginners and students with rudimentary skills overestimate their abilities while skilled students underestimate themselves. The authors of those research studies did not provide further details concerning over/underestimation, such as when it ceased or how long it took to disappear. In this study, the overestimation, especially of low achievers, had been viewed differently. For more details, here are some of participants' opinions: 'I thought that self-assessment was to show the teacher that I am able to write, but after I recognized that was for my benefit to know my troubles in writing', 'I overestimated myself because I was not up to the challenge', 'I did not accept the idea of being unable to write accurately and correct mistakes', etc. Furthermore, we have noticed that some underachievers underestimating their writing while attempting to advance, admitting that their writing was not up to the norms. They attempted to work harder in some way. One of them, for example, stated, 'I know that this is not good writing because my vocabulary is poor and sentence constructions are horrible'. Another participant added that her writing was terrible as compared to her classmates' writings and felt anxious when she saw her writing. However, she viewed it as a means to drive herself to work more. As a consequence, we may conclude that the outcomes were mixed and had nothing to do with low or high accomplishment as described in the literature.

In this study, the findings of self-assessment scoring when compared to TA were different: some low-achievers overestimated and others underestimated evaluating themselves, while some could do it properly. The same thing happened to high-achievers. Topping (2019) observes that the validity and reliability of self-assessment are a bit lower and more variable, with a propensity to over-estimation when compared to TA (Topping, 2019). The author suggest a number of factors to explain the impact over validity and reliability of self-assessment, "the ability of the learner, the amount of scaffolding, practice and feedback and the degree of advancement in the course, rather than chronological age…the nature of the product or performance assessed, the nature and clarity of the assessment criteria, and the nature of assessment instrumentation" (Topping, 2019, p. 15-16). In light of this, we found that the aforementioned factors may altogether result in concluding to various results regarding the correlation between self-assessment and TA. For example, we can see that both the nature and clarity of the evaluation criteria must be clear and understood in order to design reliable assessment instruments. Furthermore, this latter may have an impact on the whole assessment process. Another thing to consider is the learner's skill ability, which appears to be a crucial component in Topping's (2019) factors. This latter was proven to be insignificant in this study because there were high achievers overestimating their works and poor achievers objectively appraising their writings, and so on.

4.1.4 The Impact of Performing Anonymous Peer on First year EFL Degree Students' Writings

To answer the fourth research question, which intends to investigate APA performance in EFL writing assessment classroom the primary results employed for the discussion are shown in Table 4.4, indicating which data collection method was used and precisely which part (s) and item (s) was examined. The discussion focuses on students' APA performance, investigating how it occurs and assessing the whole process, whether successful and effective or not, and pointing up any barriers for future recommendations.

Table 4.4 *Results Used to Discuss the Fourth Research Question*

Construct	Data gathering tool	Data used for the discussion
Student anonymous peer assessment	Pre-study questionnaire	Section 2: Q3 Section 4: Q7 (Q7.1 & Q7.2)
	Writing rubrics	Analyzing PA done through using rubrics
	Follow-up section of the rubric in PA	Section 1 (1.1 &1.2) and Section 2
	Researcher journal report	Section 3 & 4.3
	Post-study questionnaire (Part I)	Section 3: Q1, Q2, Q3, Q4, Q5, Q6, Q7, Q8, Q9

According to the findings of this study, APA was valued since the submissions in peer writing pieces and peer criticism provided assisted students in improving their writing challenges. This not only broadened the scope of peer feedback supplied through peer assessment but also allowed them to use peer writings to improve their writing style, learn new terms, and correct grammar and language use issues. Given the importance of peer feedback, Strijbos and Wichmann (2018) believe that written and/or oral written remarks, in addition to scoring, foster communication and involvement, resulting in the collaborative nature of PA. Overall, all investigations, ranging from reviews of peer grading and peer feedbacking (Topping, 1998; Falchokiv & Goldfinch, 2000) to meta-analysis (Li, Xiong, Hunter, Guo, & Tywoniw, 2019; Double, McGrane, Hopfenbeck, 2019), have reported positive outcomes of peer assessment.

Basically, incorporating PA was shown to be beneficial in this study in terms of giving a space for students to engage and help one another improve their writing. Surprisingly, the students gave various peer feedback, including correction, advice, encouragements, improvement suggestions, and guidelines for further writing progress. For the corrective type, one of the students wrote, 'I see that the spelling of the word 'treits' is 'traits', and another one said, 'The sentence which started with 'my mother has several…a model for me' lacks parallelism in structure, so try to correct that'. For advice feedback, one student wrote, 'I advise you my classmate to revise the punctuation marks course because you have used them in a bad way. This is not to disappoint you but a reality that you have to accept for your benefit'. Another student said, 'My advice for you is to try to develop your paragraph and avoid writing short sentences'. As for encouragement feedback, one of the students said, 'I would like to congratulate you on your clear and understood content and I push you to write more and more because your writing has something special'. Another one said, 'I may be wrong to say that I really admired your writing as it is well-organized, correct, and understood. I liked your style in general'. As for suggestion feedback, one of them said, 'If I was you, I would write a topic sentence which contains more details to summarize the content of the paragraph, and shorter concluding sentence to end up the paragraph' while another one wrote, 'In your place, I would write more lengthy paragraph without fearing mistakes because if we don't do mistakes, we never learn'.

As PA offers students, both assessors and the assessees, an opportunity to learn in one domain, they also find themselves with intellectual challenges in front of a piece of work which enhances them to think deeply and stimulates their thinking process and social skills (Topping, 2019). Learning benefits that the assessor may gain can be through pondering cognitively, developing and providing feedback, and focusing in assessment (Yu, 2011). However, the assessee gains learning advantages by receiving and critically assessing feedback, determining which components to apply and why, and reflecting on the job or issue feedbacked (Li, Liu & Zhou, 2012). For the assessor, learning gains can be achieved through reflecting intellectually, formulating and giving feedback, and focusing on assessment (Yu, 2011). However, for the assessee learning benefits are achieved through receiving and evaluating critically the feedback done, deciding upon which aspects to implement and why, and reflecting on the task or the issue feedbacked (Li, Liu & Zhou, 2012). It has also been found that PA boosts student writing autonomy and self-efficacy, as well as favorable attitudes about writing (Moussaoui, 2012). It can therefore play a significant role in drawing students' attention to work with a higher feeling of duty and accountability (Topping, 1998). In light of this discussion, we found that PA contributed in reducing teacher reliance, motivated students to improve, and encouraged them to be demanding and competitive in order to face their writing challenges and begin addressing their problems.

PA was also a way for students to enhance their self-assessment in this study. To that end, the students explicitly emphasized their need for extra practice with PA, while cautioning against extending that for self-assessment to prevent boredom. In this vein, Looney (2008) sees that peer learning is important in developing solidarity within a group and offering opportunities to accelerate learning and PA because it helps students improve their own abilities in self- assessment. Based solely on experimental and quasi-experimental investigations, Li et al. (2019) and Sanchez, Atkinson, Koenka, Moshontz, and Cooper (2017)'s meta analysis revealed that PA had a favorable and substantial influence on student learning when compared to students who did not receive PA. Furthermore, Li et al. (2019) found that PA was more important than self-assessment, although the latter's implementation was quite limited leading to an unclear position. On the other hand, in Sanchez et al. (2017)'s meta-analysis, the authors discovered that self-assessment was somewhat more essential than PA; as a result, they did not rush to a definitive conclusion when comparing the two. However, Li et al. (2019) have surprisingly found that PA has more positive impacts than TA, and Harrison, O'Hara, and McNamara (2015) have concluded that assuming TA, as compared to PA, is less effectiveness when it comes to facilitating students' development toward self-reliance.

Another significant finding in this study was that students took their PA seriously, which increased discussion and involvement with their peers. The assessors had taken their peer assessors' feedback seriously, not just accepting it but also discussing it further. Here are some examples regarding that, 'If you say that my sentences are not well constructed, do you mean that all of them are not correct, or just some of them? Also, you did not say which sentence errors I did; is it fragment, choppy, non parallel structure, etc.', 'I disagree with you concerning my use of the rubric because you didn't explain where I did mistakes. So that doesn't give you the right to say that I didn't use the rubric correctly'. In addition to that, they, most of the time, thanked them for their reaction and, in some case, suggested for them how to deliver feedback next time. For example, one of them said, 'Thank you for the feedback you gave me. I really liked it', 'another added, 'I would like to thank you for your feedback, but the way you wrote it was horrible. You could use soft words instead of just blaming me'.

In terms of peer grading and over/underestimation, we unveiled that high performers tended to underestimate their peers, but they also rated them objectively. A number of poor achievers, we found that some provided correct assessments, while others undervalued or exaggerated their peers' works. This is to indicate that there is no precise result to expect. For further clarification, we must remind that the high performers in this study were largely boys, which might be attributable to a gender difference, since we saw in the low achievers category that mostly females underestimated themselves while boys did the opposite. The results are not cross checked with other studies because we found no study discussing under/overestimation in PA. In contrast, with a few exceptions, the peer feedback process was usually fair and correct. As an exception, a few low-achievers were discovered underestimating high-achievers and providing them with comments that did not reflect the reality of their work, while high achievers were discovered underestimating low achievers and providing them with disturbing comments and remarks.

4.1.5 The Impact of Assessment for Learning on First Year EFL Degree Students Writings

To address the fifth research question, which seeks to investigate the influence of AfL on EFL students' writing, the primary findings employed for discussion are shown in Table 4.5, indicating which data collection tool was used and precisely which part (s) and item (s) were examined. The discussion emphasizes the outcomes of adopting AfL, including the preceding research questions discussed (see Sections, 4.1.1, 4.1.2, 4.1.3, and 4.1.4), as well as data received from the working portfolios, researcher journal report, and post-study questionnaire responses.

Table 4.5 *Results Used to Discuss the Fifth Research Question*

Construct	Data gathering tool	Analyzed data used for the discussion
AfL and EFL writing	Post-study questionnaire (Part II)	Section 1: Q1, Q2, Q3, Q4, Q5 Section 2: Q1 (Q1.1, Q 1.2, Q1.3) Q2, Q3, Q4 (Q4.1 & Q 4.2), Q5, Q6, Q7, Q8, Q9.
	Writing working portfolio	Section 1 & 2
	Researcher journal report	Section 1 &5
	Pre- and Pot-test SPSS analysis	/

Implementing AfL helps students become more aware of the need to improve their writing, engage in improvement, and respond to the progress made. Feedback from the teacher, self, and classmates influenced the process, driving them to bridge the gap between what was now understood and what was hoped to be comprehended (Ramaprasad, 1983). For first-year EFL students, the most important aim was to understand their existing difficulties, where detecting their problems in writing and seeking for a position so that problems should not be existing anymore. They had also moved away from the need for teacher direction, realizing that they were the owners of their writing growth and that it was up to them to be proactive in order to improve. For example, one of the participants expressed explicitly that he was writing for himself rather than the teacher. Another had grasped that her mistakes were still there for her to correct, not for anybody else. To highlight the importance of feedback, Black and

Wiliam (1998) see that feedback on efforts emphasizes three important components: definitions of the intended aim, evidence regarding current position, and some comprehension of a means to bridge the gap between the two.

The students of this study had also encountered what has been argued to be among the difficulties faced in Algerian EFL writing context, which are low self-efficacy and lacking regular assessment (Moussaoui, 2012). In this context, AfL had served as a strategy to boost them for further self-regulated writing, starting from getting rid of micro-errors which destruct them as L2 writers at the expense of the macro-errors, or the content (Grabe & Kaplan, 1996). In this context, Panadero, Jonsson and Strijbos (2016), for example, have found clear connections between SRL models of self-regulation and co-regulation when implementing both self- and peer assessment. This offers the students an opportunity to compare their writings at the beginning of the intervention and at the end (see Table 3.45) where they could recognize their writing problems. To clarify, the feedback provided by the student him/herself and peers play the role of "an inherent catalyst" for all self-regulated activities because students generate an internal feedback when monitoring their engagement with tasks (Butler & Winne, 1995, p.246). On the other hand, TA and guidance also play a role in assisting students gain a given degree in autonomy to work upon their writing progress. In this regard, Wertsch (1979) proposed that regulation had stages, moving from other-regulation to self-regulation i.e. from others' assistance to being completely independent and self-reliant. This is similar to the zone of proximal development by Vygotsky (1978), where the author explains the importance of social interaction using language as a medium to grow the zones with potential for that. This is what happened with student participants as they moved from needing others' assistance to being able to face their own writing problems.

Another important result worth highlighting is that classroom assessment effectiveness in this study was essentially shaped by rubrics and student assessments, self- and peer assessment, followed by teacher feedback and guidance and working portfolio. In fact, the importance of using rubrics as assessment tools lies in helping, to a great extent, student engagement in the assessment processes. In conjunction with this, it has been concluded that if students understand what is expected of them in an assessment activity (criteria) and how well they are expected to execute it (standards), they may be able to better target their learning efforts (Armstrong, Chan, Malfroy, & Thomson, 2008). On the other hand, we may mention that teacher feedback was shown to be beneficial, but it had to be used with caution since it may reverse the desired benefits. In this context, it has been found, for example, that teacher feedback may either both silence students' voices and imposes teacher expectations, or it can empower students to produce writings that satisfy the standards in a specific environment (Hyland & Hyland, 2006).

Another important finding yielded is that there is a considerable link between self- and PA. For example, students determined whether their writings were the best or worst in PA based on self-assessment, and vice versa. The significance of this link was also demonstrated when students compared student assessment, self- and PA, in order to reach a conclusion about their actual issues, progress, and motivation. In this context, self- and PA are seen as practices of AfL which emphasize student feedback and are thought to increase student assessment involvement (Panadero, Jonsson, & Strijbos, 2016), engage student in learning overall (Ndoye, 2017), assist students build reflective and critical thinking abilities as well as self-confidence (Logan, 2009), instill in students a sense of personal responsibility for their own learning (Yorke & Longden, 2004), and commit the essential and required efforts for future learning successes (Ndoye, 2017). Furthermore, Dickhut (2003) found that this combination encourages knowledge and creativity, and higher order thinking skills. However, we should point to the fact that students in this study preferred either self- or peer assessment, and that they were more careful in PA than in self-assessment, right from the beginning of the intervention, whereas self-assessment happened incrementally after recognizing their responsibility for their writing progress.

Unpredictably to what had been believed, the results show that students played effective and interesting roles despite their various levels in writing. They liked the idea of giving their opinions before embarking on any task or stage as that gave them explicit vision to communicate their needs to teachers, alert them about their problems, and make writing classroom effective. They proved that they were an interesting source of feedback either when being involved in rubric design and classroom assessment decisions, when providing their peers with feedback, or even when assessing their high school teachers' feedback in writing classrooms. In agreement with that, Andrade and Boulay (2003) proved that students are source of feedback when given the appropriate conditions and supports. Moreover, many researchers like Bangert-Downs, Kulik, Kulik, and Morgan (1991), Kulhavy and Stock (1989), and Meyer (1986) argued that students are more effective when invited to provide feedback.

Regarding portfolio impact, in this study, students emphasized the necessity of using the working type, also known as the growth/learning portfolio, because it reminded them of the content and quality of their work, as well as the flaws they needed to correct. In this vein, Aydin (2010) argued that influence of portfolio evaluation on EFL writing was believed to have beneficial consequences despite being not enough researched. Rolheiser, Bower, and Stevahn (2000) noted that this type of portfolio incorporates.

> evidence of struggle, failure, success, and change. The growth will likely be an uneven journey of highs and lows, peaks and valleys, rather than a smooth continuum. What is significant is that learners recognize growth whenever it occurs and can discern the reasons behind that growth. The goal of a growth portfolio is for learners to see their own changes over time and, in turn, share their journey with others. (p.4)

The working portfolio in this study displayed for students their progress in thoroughly examining it; searching for missing components in their writings and how to repair them, focusing on explaining why a particular writing was done in that way, and justifying why a given assessment was provided as such.

Regarding the pre- and post-test analysis, the experiment group outperformed the control group quantitatively despite having some outliers in the data; nonetheless, these are not signalling anomalies but are attributable to large samples (Allen, 2017). This is provided for the sake of accuracy. We did not rely on quantitative data to demonstrate the importance of AfL in our study, but this is an additional factor that may reinforce and consolidate the whole discussion.

The aspects of autonomy adopted in this study are the ones by Benson (1997) technical and psychological versions which were in compatible with Cotterall's (1995) learners' behavior and beliefs. In this study, we did determine which component, either rubric, self, peer or TA, or portfolio which led to independence of the students, but we have to mention that PA was highly appreciated by students and helped them to improve their writing. On the other hand, for self-assessment, the student advised against doing it for an extended time period since it may become tedious, whereas PA required much experience. For TA, it was mostly used for direction rather than as a competitive technique for student assessment. Rubrics were significantly preferred as assessment tools and seen as a powerful instrument to assist in writing growth and advancement. As a result, we can claim that employing rubrics and APA were the primary factors that helped students advance, without discounting the relevance of TA coaching and self-assessment, as well as the working portfolio, which was viewed as a significant intervention.

4.2 Recommendations for Further Research

After discussing the findings around the research questions within the limitations' frame, in this section we present an extensive section of recommendations to guide, inspire, or encourage other researchers who are willing to invest in this field. The recommendations generated are divided into four primary areas based on the constructs under examination which are basically writing rubrics, self-assessment, PA, and AfL classroom.

4.2.1 Further Recommendations for Implementing Writing Rubrics Effectively

Rubrics are ineffective on their own; nevertheless, some precautions and requirements have to be taken into consideration. In this context, Panadero and Jonsson (2013) commented on the scarcity of studies on the factors influencing the usefulness of rubrics in fostering student learning. In this part, we will present what was done when writing rubrics were used in this study and what has to be done in the future to guarantee better usage and writing rubric implementation.

a. Designing Rubrics. It is critical to provide clear and understandable rubrics. Rubrics must have clearly specified indications and criteria in order to align both students and teachers’ expectations of the assessment activity (Cox, Morrison, & Brathwaite, 2015). To that end, we propose integrating participants in rubric construction, teaching them about rubrics, and requesting them to develop some rubrics in advance to present and discuss in class. This may cast light on how to approach rubrics, elaborate on their writing rubrics’ perceptions, and ameliorate rubric design process.

b. Implementing Rubrics. Following her experience, Crawford (2001) believes that presenting the rubric before constructing the assignment is a must since it helps both instructors and learners know what is expected of them; namely advice and establish assignment expectations. As a consequence, we recommend introducing rubrics for participants, dissociating its components and introducing each separately to compare the effects after; for example, introducing rubrics with brief descriptors but no scoring strategy, brief descriptors and a brief scoring strategy, detailed descriptors but no scoring strategy, and detailed descriptors and a detailed scoring strategy. Comparing the outcomes yielded provides insight into the best sort of rubric to create, holistic or analytic, and raises participants’ awareness of its significance. As a proposal, we prefer the analytic type above the holistic one to allow participants, especially those who did not attain advanced levels, to completely comprehend the criteria and standards of writing properly and fluently.

c. Piloting Rubrics and Training Students. Before introducing writing rubrics, the researcher highly advise testing rubrics on a regular basis to check their validity and clarity, as certain issues may not be apparent after only one or two piloting sessions. This is also true for training sessions. At last, the final version of rubrics, or properly could be called preliminary final version as some changes might be applied after training sessions, should be attempted with student participants to identify any issues with its use. As a result, we can conclude that piloting and training sessions should be accompanied since after the trial sessions, changes and revisions to the rubrics should be applied to provide the final version. In other words, the piloting phase does not always determine the final version of the rubric to be implemented; rather, the piloting and training phases may altogether collaborate to make this decision**.**

d. Time and Using Rubrics. Andrade and Boulay (2003) finished their study with some potential changes to the intervention, one of which was increasing the treatment time. At the end of the intervention we conducted, we found that time factor is true, particularly with regard to the narrative genre because the participants explicitly voiced their desire for that.

e. Gender and Rubric Use. Despite not being one of the study's objectives, we have observed some gender differences in participants' reactions to using writing rubrics and performing self- and APA. This theme has the potential to be a highly intriguing study topic. This recommendation is raised due to the considerable proportion of male participants in the experiment group. In fact, the level of self-efficacy is one of the influences highlighted in gender and using rubrics, but the results remain conflicting. For example, Andrade, Du, and Mycek (2010) found that girls had stronger self-efficacy than boys and Pajares and Valiante (1999) interpreted this as girls having more self-confidence in producing their own feedback than boys. Covill (2012), on the other hand, found no effect of rubric use on females' self-efficacy. He highlighted that students' assurance of success may lead to carelessness, but he emphasized the high self-efficacy for writing course assignments in general.

4.2.2 Further Recommendations for Effective Self-assessment Implementation

Following the discussion made of self-assessment in classroom writing assessment of the study conducted, in what follows we suggest some recommendations for future research to best approach self-assessment.

a. Performing Self-assessment Accompanied with Reflective Journals. Using reflective journals and not only a follow-up section of a rubric when performing self-assessment would be more helpful. The students would be more comfortable giving their ideas on self-assessment in the reflective journals, where they would have to describe their challenges and the benefits generated over a period of time recording it promptly. They might ponder the process and its results at any time.

b. Performing Self-assessment Using Rubrics or Other Tools. We advocate introducing self- assessment using rubrics or other similar tools against doing so without supporting this performance with a set of assessment criteria. To effectively implement self-assessment, students should have a clear understanding of the learning goals so that they may arrange their work accordingly. To that purpose, the assessment criteria, either in rubrics or scripts, should be presented with students prior to the learning processes (Panadero & Alonso-Tapia, 2013; Panadero, Jonsson & Strijbos, 2016).

Another reason to motivate researchers to enter this study field is that self-assessment using rubrics is a relatively unexplored topic (Wang, 2016). We confirmed this point by searching some databases for scholarly written research papers, such as Google Scholar and PsychNet, and finding only a few studies, such as those by Andrade, Du, and Wang (2008) and Andrade, Du, and Mycek (2010), as well as some very recent studies, such as those by Wong (2019), Salim and Suppramaniam (2020), Vasileiadou and Karadimitriou (2021), and Pui, Yuen, and Goh (2021). All of these studies did not focus on students in their first year of university, but rather on students before entering university. To clarify what we mean by conducting self-assessment using assessment tools, we refer to Chelli (2013)'s study attempting to increase EFL students' writing accuracy, grammatical complexity, and organization through self-assessment using portfolios. The findings revealed a considerable improvement in writing ability, writing attitudes, and meta-cognitive capabilities. This is an example of one of the studies that fall under this heading.

c. Self-grading and Self-assessment. When self-assessment is solely self-grading, this may fail to meet the purpose of boosting self-regulation. As a result, we strongly advise against depending purely on self-grading, and this should not be done prior to self-assessment for feedback reasons. Otherwise, grades may hinder the genuine growth desired by adopting self-assessment.

d. Time and Self-assessment. We do not propose employing self-assessment for an extended amount of time in order to avoid boring learners and preventing them from progressing because the participants of this study had clearly expressed that. However, it would be better to ask students, during the process, when they feel satisfied of their self-achievement through self-assessment to decide upon ending the practice.

e. Training and Self-assessment. We highly advise teachers and researchers to engage students in practicing self-assessment before embarking on the intervention to have enough time to identify any issues or impediments. In a word, training should take enough time to avoid unexpected complications.

f. Gender and Self-assessment. Self-assessment and gender is another issue that has been evolved in the literature. Following Gipps (1994) and Topping (2019), this difference concerns the process itself and the negotiations with teachers. While males tend to challenge teacher's assessment and having keen sense to compete for the final decision, females tend to bargain cautiously near the end. Despite not being the major interest of the study conducted, we had observed that males do not discuss too much as compared to girls.

g. Self-assessment and Final Writing Product. Instead of directing students' focus merely to self-feedback the final writings, we see that they should be given the chance to reflect on their outlines and drafts before producing their final copies, rather than keeping that only at the end of the writing process.

h. Self-assessment and Under/overestimation. When conducting studies on this theme, light might be shed on whether the results mixed, i.e. if high performers overestimated and/or underestimated their writings, and vice versa for poor achievers. Also, if high performers underestimate while poor achievers overestimate, and so forth. Furthermore, one might delve further to find the reasons behind the over/underestimation.

Because of the inconsistent results, which in particular might be an important research topic, we urge further complex investigations. It is possible to shed light on whether the results were mixed, i.e. if high achievers exaggerated and underestimated their works, and vice versa for low achievers. Also, if high achievers under-perform while low achievers over-perform, and so on. Furthermore, one might go deeper to uncover the reasons behind the over/underestimation.

4.2.3 Further Recommendations for Effective PA Implementation

Despite its importance and benefits (Sebba et al., 2008; Li, Liu, & Steckelberg 2010), PA is seen a difficult process to accomplish (Topping, 2003; 2019). Following our discussion of APA in classroom writing assessment in this study, we suggest some recommendations for future research.

a. Performing Anonymous Peer Assessment Accompanied with Reflective Journals.

Using reflective journals and not only a follow-up section of a rubric when performing PA would be more helpful. Through using reflective journals, students would give their opinions about PA more comfortably where they have to explain their difficulties and the advantages yielded while doing it over a period of time, tracking that punctually. It means they will reflect upon the process per se and its outcomes at any time instead of being limited to do that only in the classroom.

b. Anonymous/Identified Peer Assessment. We strongly recommend anonymity in PA to avoid social conflict and ensure objectivity in assessment. This does not prevent to conduct a study where to compare the effect of APA to the indentified one. For recall, anonymity was highly supported by the participants of this study. In line with this idea, several studies found that students felt uncomfortable while criticizing each other's work and found it difficult to score their peers (e.g. Topping, 1998; Falchikov & Goldfinch, 2000; Li et al., 2019; Double, McGrane & Hopfenbeck, 2019). Despite emphasizing the necessity of anonymity in providing safety to the process, Topping (2021) implies that feedback from a known source may be more impactful than feedback from an anonymous source. In this respect, it would be extremely interesting to do a study comparing identified and APA concentrating on different parts of the process such as feedback quality and style and inter-rater reliability in pure quantitative PA.

c. Performing PA Using Rubrics or Other Tools. We advocate implementing PA using rubrics or similar tools against doing so without showing assessment standards or criteria for students. According to Panadero and Alonso-Tapia (2013) and Panadero, Jonsson, and Strijbos (2016), in order to effectively implement PA, students should have a clear understanding of the learning goals in order to plan their work accordingly, sharing the assessment criteria, either in rubrics or scripts, before the learning processes.

d. Peer Grading and Peer Assessment. When PA is merely peer grading, it cannot be considered PA because the major component which is supposed to be the feedback is not supplied despite that one can consider scores as a kind of feedback. In fact, this is a dilemmatic issue because PA has progressed from peer grading to peer feedback to a mixed one; grading and feedbacking. Still, PA is supposed to be peer grading in many studies, especially at the emergence of this concept. Regardless of the opposing viewpoints, we strongly advise against depending solely on peer grading, especially when incorporating PA in AfL.

e. Time and Peer Assessment. We found that the time spent on APA was greater than the time spent on self-assessment. This disparity might also be explained by writing type. In this study, descriptive writing took less time than narrative writing. As a result, the issue of time must be considered in order to avoid either incomplete progress or blocking it entirely. We found that many students felt that they needed more time for PA than for self-assessment, particularly for the narrative genre.

f. Training and Peer Assessment. We strongly counsel students to be trained in PA before embarking on the treatment. So, it must take enough time to identify any obstacles or issues in order to provide opportunities for students to comprehend the process and its value, as well as the way in which it is carried out. In their meta-analysis, Li et al. (2019) concluded that there is a need for training to implement PA in classrooms (Sebba et al., 2008; Sanchez et al., 2017), through programs to be an integral component (Li et al., 2019), or by including it in homework activities to give students enough time to think about it. Training may comprise two sections: one on peer grading and the other on feedback delivery, with emphasis on style and content.

In this regard, we propose conducting studies comparing training students to peer grading or/and peer feedback to a control group that did not receive training. When considering whether or not to teach students to offer feedback, this might entail instructing them what to include in the feedback, i.e. the content, and how to write, i.e. the style, whereas the other group delivers comments based on their own knowledge without being constrained with a set of guidelines or instructed how to do that.

g. Gender and Peer assessment. Despite not being one of study's interests, we have observed some gender variations in PA especially in delivering peer feedback; boys tended to be brief where it was necessary and not where it was not whereas girls tried to explain and detail the feedback in all situations. Thus, this is a worthwhile research topic. As previously pointed to, gender difference was also observed in peer grading as we discussed it just above in Section 3.4.

h. Online/Compute-mediated Peer assessment. As we have only experienced paper-based PA, and given the importance of internet or computer-mediated PA in research, we strongly advocate this method of assessment, but with anonymity as a matter of suggestion. The findings concerning this aspect of the research were inconsistent. For example, it has been concluded that computer-mediated PA has a greater impact than paper-based PA because it provides greater flexibility, efficiency, and ease of access (Chen, 2016) while also assuring anonymity and random assignment (Cho & Schunn, 2006). However, computer-based PA is considered as a potential source of problems such as off-task talk (Chen, 2021). This second point may not be considered an issue because we may enable synchronized discussion, in which students communicate discussions and feedback each other at the same time, or trying offline softwares. Li et al. (2015), on the other hand, found that paper-based PA was better than computer-based PA. However, Li et al. (2019) found in their meta-analysis that computer-mediated PA may be preferable over paper-based methods.

Because none of the researchers had mentioned it, we would prefer to call this type of PA as electronic PA rather than online or computer-mediated PA. For further explanation, we see that electronic PA has a broad meaning, which means that PA is not done manually or on paper but rather through an electronic device (computer, phone, or tablet), and it may be done online or only through software that works without internet. When we say online, we imply that it is done through an electronic device and exclusively through the use of the internet. On the other hand, when we say computer-mediated, we do not always mean it needs to be done through computers; it might be done by other tools and devices, such as phones. Also, this might be accomplished without the use of the internet or simply by installing software that works offline. As a researcher, I propose that researchers who want to investigate this field describe the electronic or computer-mediated PA they are referring to because I have read many research papers where the authors did not fully describe how that online PA was gone and they were not giving due attention to the terms used and without being consistent throughout their research paper.

i. Responsiveness to Peer Assessment. We highly advise researchers to provide student participants an opportunity to respond to their peers' feedback. In our case, we did not offer that because of shortage of time. However, we strongly believe that would improve the interaction and vary feedback types and content as well as its effectiveness. This might also be viewed as a technique to provide students with equal opportunities to defend their work and address their flaws in depth. However, overextending the discussion would cause the section to lose its importance and become just a forum for solely social conflicts.

j. Assessees and Assessors Interactivity in Peer Assessment. The importance of discussing the way of receiving and responding to PA would be another important area of research. By Changing students' perceptions of the potential of getting various assessments, whether real or unreal, perceived favorable and positive or negative, constructive or destructive, and so on, this may build in them higher self-confidence to confront any PA and react accordingly. In the study conducted, we found in some cases that students felt disappointed by their classmates' PA and reluctant to some feedback. In fact, that did not lead to serious consequences like refusing to engage in the process or escaping the sessions because of that. However, that remains a precaution to avoid any psychological troubles that might result from PA.

k. Peer Assessment of Final Writing Product. Despite being used with final writing versions, PA might be used during the writing development process, such as outlining and drafting, to help students improve their final written works**.**

l. Peer Assessment and Under/overestimation. Under/overestimation in PA is another important research area that needs further investigation. We had encountered some PA that both underestimating and overestimating students' works. Up to our humble search of online databases, particularly Google Scholar, we did not spot any relevant title that addressed this topic, thus this research topic might be a significant contribution.

m. Mixed Abilities and Peer assessment. We faced students with varying skills, levels and capacities in the classroom, in spite of our efforts to triangulate those abilities by mixing high and low levels as well as high with high and low with low. We urge researchers to provide ample time to repeat the same combinations in order to optimize the effect and be far of considering the contradicting results and suggestions to conduct their studies freely without pre-determined background. In this context, Topping (2021) observes that when assessors and assesses have similar talents, a lively conversation can be expected, however in the opposite scenario, a form of supremacy and monopoly in the debate may be expected. This was found not true in the study we conducted since several underachievers attempted to enhance the debate because they grasped the writing aims seriously, appropriately, and consciously.

4.2.4 Further Recommendations for Implementing Assessment for Learning in the Writing Classroom

The fifth research question generated recommendations for further research on identifying writing genres, the interaction between self- and PA, and the implementation of a writing working portfolio.

We have wondered whether working on writing genre, beginning by descriptive genre followed by the narrative one, is the reason for moving from focusing on micro-errors to macro-errors. This indicates that in this study, student participants focused on micro-errors such as grammar and others in the descriptive genre, since it was taught first, while in the narrative genre, they largely focused on macro-errors, the content. The suggestion generated from this dilemma is whether it was really the order which impacted this shift of focus or the genre itself. In other words, working on the descriptive genre helps students focus on micro-errors better than working on the narrative genre, or studying the writing descriptive genre at the first place than the narrative one offers an opportunity for students to fix up their micro-errors and embark on the macro-errors in the subsequent genres.

Another important recommendation is studying the relation between self- and peer PA in writing. In line with this idea, Falchikov (2007) sees that self-and PA are often considered as a single assessment innovation in higher education, raising questions such as whether both should have exclusive roles or go hand in hand, when similar and what relationship between them and when and how to establish it. Some axes worth to be researched are: (1) comparing and evaluating the progress when implementing either performance alone and when introduced together, and (2) which is well worth to begin with; self- or PA and whether to use them alternately or hold each performance under practice for a given period of time before switching to the second one.

Another suggestion concerns the implementation of the working portfolio. Rather than preserving this portfolio till at the end of the procedure, this kind should be carried out during fragmented time periods, guaranteeing a formative role. This may help students hold close their issues at some stage in the writing progress rather than of maintaining that at the end of the process.

The last recommendation that would impose itself is generated from the way this investigation was conceptualized. Regarding the whole study, one can wonder why it was not undertaken taking each component separately, such as introducing solely self-assessment in the classroom for a period of time and analyzing it thoroughly. We are not against it, and we do not undermine the relevance of any study, but our conception of AfL in this study took into consideration the potential pitfalls that might appear from implementing each construct separately. For example, implementing just self-assessment has shown to result in students’ leniency as well as boredom. Our advice is to motivate researchers to undertake studies taking each construct alone i.e. to conceptualize AfL in various ways. We are quite sure that the findings they gather will have a potential impact on the overall field of AfL in Algerian, in special, and EFL context, in general.

4.3 Practical Pedagogical Implications

The researcher hopes that this study will be re-implemented in Algerian writing classes, particularly for first year students, because understanding the foundations in writing aids going on to other forms of writing, specifically research academic writing. The study's findings lead to highlight the need for a set of practical pedagogical implications whose implementations might ease the difficulties and challenges face by first year students when writing in English. These implications pertain to the writing syllabus, teachers' role, the students' position in the classroom, and the entire writing teaching staff.

4.3.1 Pedagogical Implications Regarding the Official Writing Syllabus. At the syllabus level, we propose incorporating rubrics and student assessment methods, self- and PA, not as distinct activities but as essential aspects of teaching EFL writing. It is worth bearing in mind that these components are not included in the syllabus.

In terms of incorporating rubrics, we advise using the rubrics used in this study because they effectively helped students learn the fundamentals of writing such as mechanics, language use, vocabulary, cohesion, and coherence. To recall, in addition to their clarity and the details they provide, these rubrics (see Appendix 1, 2, 3) are designed in a way to respond to the assessment classroom designed. Their effectiveness lies in posing a kind of challenge on students; they teach students to adhere to the criteria and accompany descriptions in order to generate written pieces that meet the pre-established criteria.

In light of the recent emphasis on student involvement and the outcomes yielded via implementing AfL procedures, particularly with regard to student assessment, we propose incorporating student assessment performances in writing classes to raise students' awareness of the real objectives they must meet at the end of their writing journey. In terms of self-assessment, this approach had highly motivated students in this study to examine their role and take their writing development seriously. However, over reliance on it may result in loafing and lack of dedication. APA, on the other hand, has shown that students are vital sources of feedback to assist their peers' progress. This practice did not only provide an opportunity for students to gain an understanding of their peers' writing levels, but also help them to learn how to avoid errors and emphasize their strengths. Another crucial suggestion within this suggestion is the importance of introducing APA at a larger level, between groups, because this might be much more motivating and beneficial. That could happen at least twice per semester.

4.3.2 Pedagogical Implications Regarding Teachers' Role. Advocating revisiting teachers' role since it is still characterized to be in the center of the classroom, as the owner of power, without taking students' viewpoints seriously is very crucial. Teachers should be guides and helpers, not instructors and the only decision-makers in the classroom. Despite the fact that they have a set of objectives to reach at the end, which are clearly stated in the official syllabus, teachers should reconsider the authoritative role in favor of working collaboratively with their students.

To that end, we suggest a detailed section in which the teachers' role is explicitly outlined, defined, described, and even prescribed with assessment as an inherent and extrinsic component. Doing so will certainly not undermine their abilities or initiatives but may stimulate further discussion toward increasing student involvement classrooms. This might lead to even more investment in the role of students. In fact, students have the right to achieve their goals through guidance rather than strict instructions because this latter may prevent them from being the owner of their writing progress, i.e. taking charge of and controlling their own learning to be autonomous, and this may also narrow their vision toward understanding or even achieving their objectives. Involving students in assessment may provide them with more opportunities and abilities to analyze, synthesize, and evaluate.

4.3.3 Pedagogical Implications Regarding Students' Position in the Classroom Concerning students' position, we highly advise engaging them as effective agents for generating feedback by addressing all of the items required in the classroom. This should happen once the pre-determined objectives are explained at the beginning of the year. In this vein, it was found in this study that including students in roundtable discussions, such as classroom conferencing, was quite helpful as that helped in instilling in them the necessity of appraising the content of feedback. That may direct their attention against the solely importance marks have in comparison to feedback and understand that evaluating their peers' input is essential for establishing interaction and further critical engagement. As a result, students can discuss their shared writing challenges as a group because being engaged in PA might put students in a competitive or collaborative position to act and react appropriately.

4.3.4 Pedagogical Implications Regarding the Writing Staff Members

Collaboration among the writing faculty is essential. Despite having collaborative sessions, on-site discussions were common and limited regarding which courses to include, which teaching materials to use, and what to include in exams and tests. Writing staff might consider discussing how and why to adopt assessment and search for additional effective interventions to explore students' concerns to intervene accordingly. It is worth noting that there was no discussion regarding the assessment component as an inherent part of the classroom. Colleagues are encouraged to exchange students' writings across groups in order to expose their students to a diverse range of writings in order to learn from the errors and the written pieces of their peers respecting the standards and to proceed for PA.

4.4 Chapter Summary and Conclusion

This chapter is devoted to discuss the results yielded in light of the research questions formulated at the beginning of the study and present a section of recommendations for further research, regarding rubric use, implementation of self- and PA, and overall classroom-based assessment. Before proceeding for the intervention proper, the primary step was exploring students' writing difficulties in order to design the study appropriately. Through the written pre-test, students found having problems concerning, mechanics, vocabulary, language use, cohesion and coherence. Through their answers to the pre-study questionnaire, we cross checked the realistic existence of their writing problems in addition to unveiling background rationale for that as spotting the pervasiveness of teacher centeredness in their previous learning experience at schools at different levels. Following that, we moved on to introduce rubrics as evaluation tools that were advised to help students with their assessment for learning experience. Their design was meticulous, taking into account research objectives, learning goals, and student challenges. The outcomes were intriguing and encouraging in terms of perceptive involvement and activation. Students were not only directed by an assessment instrument, but they were also involved in the assessment performance as a whole through self- and APA. Introducing self-assessment was introduced in purpose of consolidating students' relation to their writing progress to appraise their writing and react seriously to their writing progress. This performance was dealt with more or less flexibly in comparison to APA. This latter was treated more carefully and perceived very effective and fruitful because students not did they exercise assessment but also learned from their peers many aspects regarding their writing, such as how to produce more complex structures and improve their overall style.

GENERAL CONCLUSION

This study's main objective was to investigate the effectiveness of AfL in enhancing EFL writing. Prior to implementing AfL procedures to assist students in overcoming writing problems, it was felt judicious to highlight EFL writing problems in the context under study through conducting content analysis of the pre-test administered at the beginning of the year in parallel to consulting out the related literature. It has been concluded that students of first year at the department of English in the University of Algiers 2, as most of the Algerian students in their first year at university, if not all of them, do have problems in writing. Those problems were various, ranging from simple grammar errors to problems in content. As a response to that, this intervention was conducted in an attempt to solve the pre-existing problems through using AfL procedures.

The idea of intervention arose from a shift in attention from traditional teaching pedagogies and general didactics to assessment, which turned up to be a foggy area that teachers seemed to be dealing with lightly and not readily; via avoiding discussing the writing criteria and delving deeper inside what, how, and why to respect a given set of criteria. Another reason for our focus is the relatively new birth of this concept of AfL in the EFL context, with unfound studies in the Algerian setting to our knowledge. This had highly encouraged us to contribute with this research study.

This study was conducted adopting a quasi-experimental design using mixed method, for triangulation and validity sake, to gather all the necessary data to answer and discuss the research questions formulated at the outset. The data gathering tools used were namely pre-and post-tests, pre- and post-study questionnaires, writing rubrics and follow-up sections, researcher journal report, and writing working portfolio. They have been designed to fit the study in terms of the aims and objectives, validity, reliability, and practicality and gather various data that aim at documenting the overall study.

Implementing AfL procedures in the writing classroom using rubrics, self- and APA to assess their effectiveness, aimed basically at generating detailed, constructive, and elaborated feedback to trigger EFL students' writing progress. The intervention was found very fruitful and an interesting experience for students, as it gained several advantages. In reality, the basic assumption regarding the influence of assessment on learning is dual: keep students learning and ensure that they will develop if they continue to strive to learn. In other words, exploiting assessment provides students with opportunities to learn rather than force them to learn.

When it came to incorporating rubrics, this gave students the chance to tease out their writing pieces and analyze them against the writing criteria, which encouraged them to work harder and look for developing and enhancing their writings even more. The students attempted to track their progress over a continuum established by the descriptors. By doing so, they learned that their challenges were well worth consideration, despite the fact that they were initially unaware of them. Basically, the assessment was considered mixed, qualitative and quantitative, and the synergy, i.e. mutual collaboration, was regarded to be fairly successful. Regarding the genres planned by the writing staff to deal with descriptive and narrative, the participants demonstrated much improvement on both genres but did spend less time while writing in descriptive genre against the narrative one.

In terms of introducing self-assessment, the participants realized that they are accountable for their own writing in general, as well as their development and difficulties. They understood their responsibility toward writing for themselves and their own advantage, not for the teacher, because they explicitly expressed that, and some of them discovered that their achievement helped them successfully write in subsequent courses, particularly in exams. It is worth noting, however, that the participants warned against completing self-assessment for an extended amount of time since it may be monotonous and therefore boring or even loafing.

On the other hand, implementing APA has been found much more motivating and challenging than self-assessment. This was due to benefiting from peers' feedback and learning from their written pieces, either good or poor. Writing samples, the good ones, taught students to improve their writing and take advantage of some aspects such as how to use grammar correctly, express effectively their idea, organize paragraphs, and so on. However, the poor samples warned students to work on their writings seriously and in the same time made them aware of the necessity of not taking their mistakes stressfully. Furthermore, participants regarded the feedback as diverse and useful. While peer feedback was various, it was regarded to be insightful, descriptive, prescriptive, corrective, and encouraging.

REFERENCES

AbiSamra, N. (2003). An analysis of errors in Arabic speakers' English writings: Investigating writing problems among Palestinian students studying English as a foreign language [Doctoral dissertation].

Adachi, C., Tai, J., & Dawson, P. (2018). A framework for designing, implementing, communicating and researching peer assessment. *Higher Education Research & Development, 37*(3), 453-467. http://dx.doi.org/10.1080/07294360.2017.1405913

Adams, P. (2006). Exploring social constructivism: Theories and practicalities. *Education, 34*(3), 243-257. http://dx.doi.org/10.1080/03004270600898893

Adams, T. L. (1998). Alternative assessment in elementary school mathematics. *Childhood Education, 74*(4), 220-224.

Ahmed, A. (2010). *The EFL essay writing difficulties of Egyptian student teachers of English: Implications for essay writing curriculum and instruction.* [Doctoral Thesis]. University of Exeter

Allal, L. (2010). Assessment and the regulation of learning. In P. Peterson, E. Baker, & B. McGaw (Eds.), *International encyclopedia of education* (3rd ed.) (pp. 348–352). Elsevier.

Allal, L. (2011). *Assessment for learning culture in the classroom, and beyond. Paper presented at the 4th Assessment for Learning conference*. Bergen.

Allen, L. Q. (2004). Implementing a culture portfolio project within a constructivist paradigm. *Foreign Language Annals, 37*(2), 232-239. http://dx.doi.org/10.1111/j.1944-9720.2004.tb02196.x

Amhag, L. (2013). Creativity in and between collaborative peer assessment processes in higher distance education. *Creative Education*, *4*(7A), 94-104.

Andrade, H. (2018). Feedback in the context of self-assessment. In A. A. Lipnevich & J. K. Smith (Eds.), *The Cambridge handbook of instructional feedback* (pp. 376-408): Cambridge University Press.

Andrade, H. G. (1997). Understanding Rubrics by Heidi Goodrich Andrade. *Educational Leadership 54*(4), 1-7

Andrade, H. G. (2000). Using rubrics to promote thinking and learning. *Educational Leadership, 57*(5), 13-19.

Andrade, H. G. (2005). Teaching with rubrics: The good, the bad, and the ugly. *College Teaching, 53*(1), 27-31. http://www.jstor.org/stable/27559213

Andrade, H. G., & Boulay, B. A. (2003). Role of rubric-referenced self-assessment in learning to write. *The Journal of Educational Research, 97*(1), 21-30. http://dx.doi.org/10.1080/00220670309596625

Andrade, H. L. (2006). Letters to EJ: The troubles with the narrow view of rubrics.

English Journal, *95*(6), 9.

Andrade, H. L. (2010). Students as the definitive source of formative assessment. In Andrade, H., & Cizek, G. J. (Eds.). *Handbook of formative assessment*, (pp. 90-105). Routledge.

Andrade, H. L., & Heritage, M. (2018). *Using formative assessment to enhance learning, achievement, and academic self-regulation*. Routledge.

Andrade, H. L., Du, Y., & Mycek, K. (2010). Rubric‐ referenced self‐ assessment and middle school students' writing. *Assessment in Education: Principles, Policy & Practice*, *17*(2), 199-214. http://dx.doi.org/10.1080/09695941003696172

Andrade, H. L., Du, Y., & Wang, X. (2008). Putting rubrics to the test: The effect of a model, criteria generation, and rubric-referenced self-assessment on elementary school students' writing. *Educational Measurement: Issues and Practice*, *27*(2), 3-13. https://doi.org/10.1111/j.1745-3992.2008.00118.x

Andrade, H., & Du, Y. (2005). Student perspectives on rubric-referenced assessment. *Practical Assessment, Research, and Evaluation*, *10*(1), 3. http://scholarsarchive.library.albany.edu/edpsych_fac_scholar/2

Andrade, H., & Du, Y. (2007). Student responses to criteria-referenced self-assessment. *Assessment & Evaluation in Higher Education*, *32*(2), 159-181. http://dx.doi.org/10.1080/02602930600801928

Andrade, H., & Valtcheva, A. (2009). Promoting learning and achievement through self-assessment. *Theory into Practice*, *48*(1), 12-19. http://dx.doi.org/10.1080/00405840802577544

Andrade, M. C., & Evans, N. (2013). *Principles and practices for response in second language writing: Developing self-regulated learners.* Routledge.

Andrews, R. (2001). *Teaching and learning English: A Guide to recent research and its implications.* Continuum International Publishing Group.

Armstrong, S., Chan, S., Malfroy, J., & Thomson, R. (2008). Assessment guide: Implementing criteria and standards-based assessment. University of Western Sydney. http://www.uws.edu.au/tdu

Arnaudet, M. & Barret, M. (1981). *Paragraph development: A Guide for students of English as a second language*. Engelwood Cliffs.

Arter, J. A., & Chappuis, J. (2007). *Creating & recognizing quality rubrics.* Pearson Merrill Prentice Hall.

Arter, J. A., & Spandel, V. (1992). Using portfolios of student work in instruction and assessment. *Educational Measurement: Issues and Practice*, *11*(1), 36-44. http://dx.doi.org/10.1111/j.1745-3992.1992.tb00230.x

Ashworth, A. (2015). *Veto on the use of null hypothesis testing and p intervals: right or wrong.* Resources Online.

Assessment Reform Group (2002). *Assessment for Learning: 10 Principles*.

Cambridge University.

Association of American Colleges & Universities (2018). *VALUE Rubric development project.* Association of American Colleges & Universities

Atkins, M. J., Beattie, J., & Dockerell, W. B. (1993). Assessment issues in higher education. Department of Employment. https://eric.ed.gov/?id=ED369370

Ausubel, D. P. (1963). *The psychology of meaningful verbal learning.* Grune & Stratton.

Aydin, S. (2010). EFL writers' perceptions of portfolio keeping. *Assessing Writing, 15*(3), 194-203. http://dx.doi.org/10.1016/j.asw.2010.08.001

Badger, R. & White, G. (2000). A process genre approach to teaching writing. *ELTJ, 54* (2), 153-160.

Bahloul, M. (2007). Spelling errors of Arab learners: Evidence of intergraphic mapping. *Language teacher research in the Middle East,* 41-51.

Baleghizadeh, S., & Dadashi, M. (2011). The effect of direct and indirect corrective feedback on students' spelling errors. *Profile Issues in Teachers Professional Development, 13*(1), 129-137.

Ballantyne, R., Hughes, K., & Mylonas, A. (2002). Developing procedures for implementing peer assessment in large classes using an action research process. *Assessment & Evaluation in Higher Education, 27*(5), 427-441. http://dx.doi.org/10.1080/0260293022000009302

Bangert-Drowns, R. L., Kulik, C. C., Kulik, J. A., & Morgan, M. T. (1991). The instruction effect of feedback in test-like events. *Review of Educational Research, 61*(2), 213–238. http://dx.doi.org/10.3102/00346543061002213

Barnett, R. (2007). Assessment in Higher Education: An impossible mission? In Boud, D. & Falchikov, N. (Eds.), *Rethinking assessment in higher education learning for the longer term*. Routledge.

Beaufort, A. (2007). *College writing and beyond: A new framework for university writing instruction.* Logan.

Becker, A. (2016). Student-generated scoring rubrics: Examining their formative value for improving ESL students' writing performance. *Assessing Writing, 29*, 15-24. http://dx.doi.org/10.1016/j.asw.2016.05.002

Belachew, M., Getinet, M., & Gashaye, A. (2015). Perception and practice of self-assessment in EFL writing classrooms. *Journal of Languages and Culture, 6*(1), 1-8.

Belanoff, P., & Dickson, M. (1991). Portfolios: Process and product. Boyton/Cook,.

Ben-David, M. F. (1999). Outcome-based education: Part 3-Assessment in outcome-based education. *Medical Teacher, 21*(1), 23-25. http://dx.doi.org/10.1080/01421599979987

Bennett, Randy Elliot (2011). Formative assessment: a critical review. *Assessment in Education: Principles, Policy & Practice, 18*(1), 5–25. https://doi.org/10.1080/0969594X.2010.513678

Benson, P. (1997). The philosophy and politics of learner autonomy. In Benson, P., & Voller, P (Eds.) *Autonomy and independence in language learning* (pp. 18-34). Routledge.

Bereiter, C., & Scardamalia, M. (2013). *The psychology of written composition*. Routledge.

Bergman, M. M. (Ed.). (2008). *Advances in mixed methods research: Theories and applications.* Sage publications.

Bernstein, B. (2000). *Pedagogy, symbolic control, and identity: Theory, research, critique* (Revised Ed.). Rowman & Littlefield.

Bethan, M. (2002). Thinking through Assessment: An Interview with Dylan Wiliam. *English Education, 36*(3), 47-60.

Bhatia, A. T. (1974). An error analysis of students' compositions. IRAL: International *Review of Applied Linguistics in Language Teaching, 12*(4), 337.

Biggs, J. (1998). Assessment and Classroom Learning: a role for summative assessment? *Assessment in Education, 5*(1), 103-110. https://doi.org/10.1080/0969595980050106

Birenbaum, M., DeLuca, C., Earl, L., Heritage, M., Klenowski, V., Smith, K., et al. (2015). International trends in the implementation of assessment for learning: Implications for policy and practice. *Policy Futures in Education, 13*(1), 117–140. http://dx.doi.org/10.1177/1478210314566733

Birjandi, P., & Hadidi Tamjid, N. (2012). The role of self-, peer and teacher assessment in promoting Iranian EFL learners' writing performance. *Assessment & Evaluation in Higher Education, 37*(5), 513-533. http://dx.doi.org/10.1080/02602938.2010.549204

Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of Second Language Writing, 17*(2), 102-118. http://dx.doi.org/10.1016/j.jslw.2007.11.004

Black, P. & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability, 21*(1), 5-31. https://doi.org/10.1007/s11092-008-9068-5

Black, P. and Wiliam, D. (1998b). *Inside the black box: Raising standards through classroom assessment*. Phi Delta Kappa.

Black, P., & Wiliam, D. (2003). In praise of educational research: formative assessment. *British Educational Research Journal, 29*(5), 623–637. https://doi.org/10.1080/0141192032000133721

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. http://dx.doi.org/10.1007/s11092-008-9068-5

Black, P., Harrison, C., & Lee, C. (2003). *Assessment for learning: Putting it into practice.* McGraw-Hill Education.

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. http://dx.doi.org/10.1177/003172170408600105

Blakeslee, A. (2001). *Interacting with audiences: Social influences on the production of scientific writing.* Mahwah, NJ: Erlbaum.

Bloom, B. S. (1968). Learning for Mastery. Instruction and Curriculum. Regional Education Laboratory for the Carolinas and Virginia, Topical Papers and Reprints, Number 1. *Evaluation Comment, 1*(2), 1-12.

Bloom, B. S. (1976). *Human Characteristics and School Learning.* McGraw-Hill.

Bloom, B., Hastings, J. T., & Madaus, G. F. (1971). *Handbook on Formative and Summative Evaluation of Student Learning.* McGraw-Hill.

Blumner, J. S., & Fritz, F. (1997). Students Using Evaluation in Their Writing Process. In Tchudi, S. (Ed.), *Alternatives to grading student writing* (pp. 233-243). National Council of Teachers of English.

Bonner, S. M. (2013). *Validity in classroom assessment: Purposes, properties, and principles.* Sage Publications.

Bordens, K. S., & Abbott, B. B. (2018). *Research design and methods: A process approach* (10th ed.). McGraw-Hill.

Borkan, J. M. (2004). Mixed Methods Studies: A Foundation for Primary Care Research. *The Annals of Family Medicine, 2*(1), 4–6. http://dx.doi.org/10.1370/afm.111

Bostock, S. (2000). Student peer assessment. *Learning Technology, 5*(1), 245-249.

Boud, D. (1999). Avoiding the traps: Seeking good practice in the use of self assessment and reflection in professional courses. *Social Work Education, 18*(2), 121-132. http://dx.doi.org/10.1080/02615479911220131

Boud, D. (2000). Sustainable assessment: rethinking assessment for the learning society. *Studies in Continuing Education, 22*(2), 151–167. http://dx.doi.org/10.1080/713695728

Boud, D., & Associates. (2010).Assessment 2020: Seven propositions for assessment reform in higher education. https://www.uts.edu.au/sites/default/files/Assessment-2020_propositions_final.pdf

Boud, D., & Falchikov, N. (1989). Quantitative studies of student self-assessment in higher-education: A critical analysis of findings. *Higher Education, 18*(5), 529-

549. http://dx.doi.org/10.1007/BF00138746

Boud, D., & Falchikov, N. (2007). *Rethinking assessment in higher education learning for the longer term.* Routledge.

Boud, D., Cohen, R., & Sampson, J. (1999). Peer learning and assessment. *Assessment & Evaluation in Higher Education, 24*(4), 413-426. http://dx.doi.org/10.1080/0260293990240405

Boud, D.J. (1991). *Implementing student self-assessment. In Higher Education Research and Development Society of Australia green guide*. Higher Education Research and Development Society of Australia.

Boud, D.J. (1995). *Enhancing learning through self assessment*. Kogan Page

Bouhadiba, F. (2000). ELT and cross cultural communication. *Imago, 3,* 95–110.

Bouyakoub, N. (2012). *Exploring writing in an academic context the case of first-year university students* [Doctoral Thesis]. University of Tlemcen.

Boyd, E., Green, A., Hopfenbeck, T. N., & Stobart, G. (2019). *Effective feedback: The key to successful assessment for learning.* Oxford University Press.

Boyd, E., Green, A., Hopfenbeck, T. N., & Stobart, G. (2019). Effective feedback: The key to successful assessment for learning. Oxford University Press.

Boynton, P. M., & Greenhalgh, T. (2004). Selecting, designing, and developing your questionnaire. *Bmj, 328*(7451), 1312-1315. http://dx.doi.org/10.1136/bmj.328.7451.1312

Bransford, J.D., Brown, A.L., & Cocking, R.R. (2000). *How people learn: Brain, mind, experience and school.* National Academy Press.

Braun, V. & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101. http://dx.doi.org/10.1191/1478088706qp063oa

Brindley, G. (2001). Language assessment and professional development. *Experimenting with uncertainty: Essays in honour of Alan Davies, 11*, 137-143.

Broadfoot, P., James, M., Mcmeeking, S., Nuttall, D., & Stierer, S. (1988). *Records of Achievement: Report of the National Evaluation of Pilot Schemes (PRAISE)*. HMSO.

Brookhart, S. M. (2005). Assessment theory for college classrooms. *New directions for teaching and learning*, (100), 5-14. http://dx.doi.org/10.1002/tl.165

Brookhart, S. M. (2013). *How to create and use rubrics for formative assessment and grading.* ASCD.

Brookhart, S. M. (2018). Appropriate criteria: key to effective rubrics. *Frontiers in Education 3*(22). http://dx.doi.org/10.3389/feduc.2018.00022

Brooks, V. (2002). *Assessment in secondary schools: The new teacher's guide to* *monitoring, assessment, recording, reporting and accountability*. Open University Press.

Brown, G. T. L., & Harris, L. R. (2013). Student self-assessment. In J. McMillan (Ed.), *The SAGE handbook of research on classroom assessment* (pp. 367-393). Thousand Oaks, CA: SAGE.

Brown, S. (1999). Institutional strategies for assessment. In Brown, S., & Glasner, A. (Eds.), *Assessment matters in higher education: Choosing and Using Diverse Approaches.* SRHE and Open University Press

Bryman, A. (2012). Social research methods (4th ed.). Oxford University Press.

Burner, T. (2014). The potential formative benefits of portfolio assessment in second and foreign language writing contexts: A review of the literature. *Studies in Educational Evaluation*, 43, 139-149. http://dx.doi.org/10.1016/j.stueduc.2014.03.002

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, *65*(3), 245-281. http://dx.doi.org/10.3102/00346543065003245

Calkins, L., Ehrenworth, M., & Akhmedjanova, D. (2019). Creating Formative Assessments Systems in the teaching of writing, and harnessing them as professional development .In Andrade, H., L., Bennett, R., E., & Cizek, G. (Eds.), *Handbook of formative assessment in the disciplines* (pp. 173-206). Routledge.

Cameron, C., Tate, B., Macnaughton, D., & Politano, C. (1998). *Recognition without Rewards*. Peguis Publishers

Campbell, D. M., Cignetti, P. B., Melenyzer, B. J., Nettles, D. H., Wyman Jr, & R. otros (1997). *How to develop a professional portfolio.* Pearson Higher Ed.

Campbell, D. T. (1957). Factors relevant to the validity of experiments in social settings. *Psychological Bulletin, 54*(4), 297.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*(2), 81.

Campbell, D.T.& J.C. Stanley(1966). *Experimental and Quasi-experimental Designs for Research.* Rand McNally.

Canagarajah, S., & Jerskey, M. (2009). Meeting the needs of advanced multilingual writers. In Beard, R., Riley, J., Myhill, D., & Nystrand, M. (Eds.), *The SAGE handbook of writing development* (pp. 472-488). Sage Publications.

Careless, D., (2017). Scaling Up Assessment for Learning: Progress and Prospects. In David Carless, D., Bridges, S., M., Chan, C., K, & Glofcheski, R. (Eds.), *Scaling up assessment for learning in higher education* (pp. 3-17). Springer

Carless, D. (2007). Learning-oriented assessment: conceptual bases and practical implications. *Innovations in Education and Teaching International*, *44*(1), 57-66. http://dx.doi.org/10.1080/14703290601081332

Carnicelli, T. (1980). The writing conference: A one-to-one conversation. In T. Donovan & B. McClelland (Eds.), *Eight approaches to teaching composition* (pp. 101-131). IL: National Council of Teachers of English.

Chakraverty, A., & Gautum, K. (2000). Dynamics of writing. *Forum, 38*(3), 22-25.

Chase, C. I. (1999). *Contemporary assessment for educators*. Addison-Wesley Longman.

Chelli, S. (2012). The effects of the competency based approach on leaner's writing achievement a case study of first-year students at Biskra university [Doctoral Disseration]. University of Biskra.

Chen, J. (2021). The effectiveness of peer assessment in EFL blended learning. *Computers, information processing and advanced education* 566-568. http://dx.doi.org/10.1145/3456887.3457015

Cheng, W., & Warren, M. (2005). Peer assessment of language proficiency. *Language Testing*, *22*(1), 93-121. http://dx.doi.org/10.1191/0265532205lt298oa

Chickering, A. W., & Gamson, Z. F. (1987). *Seven principles to good practice in undergraduate education.* The Johnson Foundation Inc.

Childers, L., & Kent, K. (2007). *Guidelines for writing a thesis or dissertation.* University of Florida.

Cho, K., Schunn, C. D., & Charney, D. (2006). Commenting on writing: Typology and perceived helpfulness of comments from novice peer reviewers and subject matter experts. *Written Communication*, *23*(3), 260-294. http://dx.doi.org/10.1177/0741088306289261

Cho, K., Schunn, C. D., and Wilson, R. W. (2006). Validity and reliability of scaffolded peer assessment of writing from instructor and student perspectives. *Educational Psychology, 98,* 891–901. http://dx.doi.org/10.1037/0022-0663.98.4.891

Christensen, L., Johnson, B, Turner, L, et al. (2014). *Research methods, design, and analysis* (12th edition). Pearson Education

Clark, I. (2012). Formative assessment: Assessment is for self-regulated learning. Educational *Psychology Review*, *24*(2), 205-249. https://doi.org/10.1007/s10648-011-9191-6

Coe, R. E., & Freedman, A. (1998). Generative rhetoric. *Theorizing Composition: A Critical Sourcebook of Theory and Scholarship in Contemporary Composition Studies*, 131-136.

Coe, R. M., & Freedman, A. (1998). Genre theory: Australian and north American approaches. *Theorizing composition: A critical sourcebook of theory and scholarship in contemporary composition studies*, 136-147.

Coffin, C., & Donohue, J. P. (2012). Academic Literacies and systemic functional linguistics: How do they relate? *Journal of English for Academic Purposes*, *11*(1), 64-75. http://dx.doi.org/10.1016/j.jeap.2011.11.004

Cohen, A., & Cavalcanti, M. (1990). Feedback on compositions: Teacher and student verbal reports. In. B. Kroll (Ed.), *Second language writing: Research insights for the classroom* (pp. 155–177). Cambridge University Press.

Cohen, L., Manion, L. & Morrison, K. (2007). *Research methods in education* (6th ed.). Routledge.

Colbert, P., & Cumming, J. J. (2014). Enabling all students to learn through assessment. In Wyatt-Smith, C., Klenowski, V., & Colbert, P. (Eds.), *Designing assessment for quality learning* (pp. 211-231). Springer, Dordrecht.

Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *The Journal of the Learning Sciences*, *13*(1), 15-42. http://dx.doi.org/10.1207/s15327809jls1301_2

Connor, U. (1996). *Contrastive rhetoric: Cross-cultural aspects of second-language writing*. Cambridge University Press.

Cook, T. D., Campbell, D. T., & Day, A. (1979). *Quasi-experimentation: Design & analysis issues for field settings*. Houghton Mifflin.

Cotterall, S. (1995). Readiness for autonomy: Investigating learner beliefs. *System*, *23*(2), 195-205. http://dx.doi.org/10.1016/0346-251x(95)00008-8

Coulmas, F. (2003). *Writing systems: An introduction to their linguistic analysis*. Cambridge University Press.

Council, B. (2018). *IELTS Task 2 Writing Band Descriptors: Public Version*. British Council.

Covill, A. E. (2012). College students' use of a writing rubric: effect on quality of writing, self-efficacy, and writing practices. *Journal of Writing Assessment*, *5*(1).

Cowan, J. (1988). Struggling with student self-assessment. In D. Boud (Ed.). *Developing student autonomy in learning* (2nd Ed.) (pp, 192-210). Kogan Page

Cox, G., Morrison, J., & Brathwaite, B. (2015). The rubric: an assessment tool to guide students and markers. In 1st International Conference On Higher Education Advances (Head'15) (pp. 26-32). Editorial Universitat Politècnica de València.

Crawford, C. M. (2001). Rubrics: Models of evaluation within a constructivist learning environment.

Creswell, J. W. (1999). Mixed method research: Introduction and application. In T. Cijek (Ed.), *Handbook of educational policy* (pp. 455–472). Academic Press

Creswell, J. W. (2009). Mapping the field of mixed methods research. *Journal of Mixed Method Research*. *3*(2), 95-108. http://dx.doi.org/10.1177/1558689808330883

Creswell, J. W. (2014). *Research design: Quantitative, qualitative and mixed methods* (4th edition). Sage Publication.

Creswell, J. W., Klassen, A. C., Plano Clark, V. L., & Smith, K. C. (2011). Best practices for mixed methods research in the health sciences. *Bethesda (Maryland): National Institutes of Health*, 541-545.

Creswell, J. W., Plano Clark, V. L., Gutmann, M. L., & Hanson, W. E. (2003). Advanced mixed methods research designs. In Tashakkori, A & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 209–240).Thousand Oaks, CA

Crookes, G., & Chaudron, C. (1991). Guidelines for classroom language teaching. In Celce-Murcia, M., & McIntosh, L. (Eds.). *Teaching English as a second or foreign language* (pp. 46-67). University of California, Los Angeles.

Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research*, *58*(4), 438-481. http://dx.doi.org/10.3102/00346543058004438

Crossouard, B., & Pryor, J. (2012). How theory matters: Formative assessment theory and practices and their different relations to education. *Studies in Philosophy and Education: An International Journal*, *31*(3), 251-263. https://doi.org/10.1007/s11217-012-9296-5

Cumming, A. (1989). Writing expertise and second language proficiency. *Language Learning*, *39*(1), 81-135. http://dx.doi.org/10.1111/j.1467-1770.1989.tb00592.x

Cumming, A., & Riazi, A. (2000). Building models of adult second-language writing instruction. *Learning and Instruction*, *10*(1), 55-71. http://dx.doi.org/10.1016/s0959-4752(99)00018-3

Cumming, G. (2013). The new statistics: Why and how. *Psychological Science*, *25*(1), 7-29. http://dx.doi.org/10.1177/0956797613504966

Sadler, R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*(2), 119–144. https://doi.org/10.1007/bf00117714

Danielson, C., & Abrutyn, L. (1997). *An introduction to using portfolios in the classroom*. Association for Supervision and Curriculum Development

Das, B. K. (1978). An Investigation of some advanced skills of composition. *CIEFL Bulletin*, *14*(1), 43-62.

Davies, A., & LeMahieu, P. (2003). Assessment for learning: Reconsidering portfolios and research evidence. In Segers, M., Dochy, F., & Cascallar, E. (Eds), *Optimising new modes of assessment: In search of qualities and standards* (pp. 141-169). Springer, Dordrecht.

DeCosta, M., & Roen, D. (2015). Assessing the teaching of writing: A scholarly approach. In Dayton, A. E. (Ed.), *Assessing the teaching of writing: Twenty-first century trends and technologies* pp (13-30). University Press of Colorado.

DeMarrais, K. B., & Lapan, S. D. (2004). Introduction. In DeMarrais, K. B., & Lapan, S. D. (Eds.), *Foundations for research: Methods of inquiry in education and the* *social sciences* (pp.1-12). Routledge.

Denzin, N. K. (1970). Problems in analyzing elements of mass culture: Notes on the popular song and other artistic productions. *American Journal of Sociology*, *75*(6), 1035-1038. http://dx.doi.org/10.1086/224853

Denzin, N. K., & Lincoln, Y. S. (2008). Introduction: The discipline and practice of qualitative research. In N. K. Denzin, & Y. S. Lincoln (Eds.), *Strategies of qualitative inquiry* (pp. 1–43). Sage Publications, Inc.

Derakhshan, A., & Karimi, E. (2015). The interference of first language and second language acquisition. *Theory and Practice in Language Studies*, *5*(10), 2112.

Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to the educative process*. Lexington, Mass.:Heath.

Diab, N. (1996). The transfer of Arabic in the English writings of Lebanese students'.

Diab, R. L., (2006). Teaching practices and student learning in the introductory research methods class. *Teaching English as a Second or Foreign Language 10*(2), 1-26. http://www.tesl-ej.org/ej38/a6.html

Diab, R., & Balaa, L. (2011). Developing detailed rubrics for assessing critique writing: Impact on EFL university students' performance and attitudes. *TESOL Journal*, *2*(1), 52-72. http://dx.doi.org/10.5054/tj.2011.244132

Dickhut, J. E. (2003). A brief review of creativity. www.personalityresearch.org/papers/dickhut.html

Diederich, P. B., French, J. W., & Carlton, S. T. (1961). Factors in judgments of writing ability. *ETS Research Bulletin Series*, *1961*(2), i-93. http://dx.doi.org/10.1002/j.2333-8504.1961.tb00286.x

Dörnyei, Z., & Taguchi, T. (2010). *Questionnaires in second language research: Construction, administration, and processing* (2nd ed.). Routledge.

Double, K. S., McGrane, J. A., & Hopfenbeck, T. N. (2020). The impact of peer assessment on academic performance: A meta-analysis of control group studies. *Educational Psychology Review*, *32*(2), 481-509. http://dx.doi.org/10.1007/s10648-019-09510-3

Draper, S. (2007). A Momentary Review of Assessment Principles. The REAP International Online Conference on Assessment Design for Learner Responsibility. http://www.reap.ac.uk/reap/reap07/Portals/2/CSL/keynotes/david%20nicol/A_momentary_review_of_assessment_principles.pdf

Drever, E. (1995). Using semi-structured interviews in small-scale research. The Scottish Council for Research in Education.

Duncan, M. (2007). Whatever happened to the paragraph? *College English*, *59*(5), 470-495. http://dx.doi.org/10.2307/25472231

Earl, L. M. (2003). Assessment as learning: Using classroom assessment to maximize student learning. Corwin Press, Inc.

Earp, B. D., & Trafimow, D. (2015). Replication, falsification, and the crisis of confidence in social psychology. *Frontiers in psychology*, *6*, 621. http://dx.doi.org/10.3389/fpsyg.2015.00621

East, M. (2008). *Dictionary use in foreign language writing exams: Impact and implications*. John Benjamins Publishing.

Educational Testing Service. (2014). TOEFL iBT® test: Independent writing rubrics.

Edwards, R., & Sutton, A. (1992). A practical approach to student centred learning. *British Journal of Educational Technology*, *23*(1), 4-20. http://dx.doi.org/10.1111/j.1467-8535.1992.tb00305.x

Efron, S. E., & Ravid, R. (2019). *Writing the literature review: A practical guide*. The Guilford Press

Elbow, P. (1973). *Writing without teachers*. Macmillan education

Elbow, P. (1990). Forward: About personal expressive academic writing. *Pre-Text: A Journal of Rhetorical Theory*, *11*, 7-20

Elbow, P. (1998). *Writing with power: Techniques for mastering the writing process*. Oxford University Press.

Ellis, R. (2015). *Understanding second language acquisition* (2nd Ed.). Oxford University Press.

Elton, L., & Jonhston, B. (2002) Assessment in Universities: a critical review of research. *LTSN Generic Centre*. http://eprints.soton.ac.uk/59244/1/59244.pdf

Enneifer, S. (2021). Tunisian EFL students' perceptions toward their writing difficulties. *Open Journal of Modern Linguistics*, *11*(4), 655-667.

Faigley, L. (1986). Competing theories of process: A critique and a proposal. *College English*, *48*(6), 527-542.

Falchikov, N. (1986). Product comparisons and process benefits of collaborative peer group and self assessments. *Assessment and Evaluation in Higher Education*, *11*(2), 146-166.

Falchikov, N. (2005). *Improving assessment through student involvement: Practical solutions for aiding learning in higher and further education*. Routledge.

Falchikov, N. (2007). The place of peers in learning and assessment. In Falchikov, N. & Boud, D.,(eds.), *Rethinking assessment in higher education: Learning for the longer term*, (pp, 128-143). Routledge.

Falchikov, N., & Boud, D. J. (1989). Student self assessment in higher education: A meta-analysis. *Review of Educational Research*, *59*(4), 395-430. http://dx.doi.org/10.3102/00346543059004395

Falchikov, N., & Goldfinch, J. (2000). Student peer assessment in higher education: A meta-analysis comparing peer and teacher marks. *Review of Educational Research*, *70*(3), 287-322. http://dx.doi.org/10.3102/00346543070003287

Fautley, M., & Savage, J. (2008). *Assessment for learning and teaching in secondary schools*. Learning Matters.

Ferris, D. R. (2002). *Treatment of error in second language writing classes*. University of Michigan Press.

Ferris, D. R. (2003). *Response to student writing: Implications for second language students*. Routledge.

Ferris, D. R., & Hedgcock, J. S. (2014). *Teaching L2 composition: Purpose, process, and practice*. Routledge.

Ferris, D., & Roberts, B. (2001). Error feedback in L2 writing classes. How explicit does it need to be? *Journal of Second Language Writing, 10*(3), 161-184. http://dx.doi.org/10.1016/s1060-3743(01)00039-x

Ferris, D.R., Pezone, S., Tade, C.R., Tinti, S. (1997). Teacher commentary on student writing: descriptions and implications. *Journal of Second Language Writing* 6(2), 155–182. http://dx.doi.org/10.1016/s1060-3743(97)90032-1

Feuer, M. J., Towne, L., & Shavelson, R. J. (2002). Scientific culture and educational research. *Educational Researcher, 31*, 4–14.

Flick, U. (1992). Triangulation revisited: Strategy of validation or alternative? *Journal for the Theory of Social Behaviour*.

Flick, U. (2018). *An introduction to qualitative research*. Sage Publications

Fontana, D., & Fernandes, M. (1994). Improvements in mathematics performance as a consequence of self-assessment in Portuguese primary school pupils. *British Journal of Educational Psychology, 64*(3), 407-417. http://dx.doi.org/10.1111/j.2044-8279.1994.tb01112.x

Friedlander, A. (1990). Composing in English: Effects of a first language on writing in English as a second language. In Kroll, B. (ed.), *Second language writing*. http://dx.doi.org/10.1017/CBO9781139524551.012

Frodesen, J., & Holten, C. (2003). Grammar and the ESL writing class. In B. Kroll (Ed.), *Exploring the dynamics of second language writing* (pp. 141-161). Cambridge University Press.

Fujeda, Y. (2006). A Brief Historical Sketch of Second Language Writing Studies: A Retrospective. *共愛学園前橋国際大学論集*, 6, 59-73

Gardner, H. (2006). *Changing minds: The art and science of changing our own and other peoples minds*. Harvard Business Review Press.

Gass, S. M., & Mackey, A. (2005). *Second language research*. Lawrence Erlbaum Associate Publishers.

Genesee, F., Upshur, J. A., & Richards, J. C. (1996). *Classroom-based evaluation in second language education*. Cambridge University Press.

Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C., & Innocenti, M. S. (2005). Quality indicators for group experimental and quasi-experimental research in special education. *Exceptional Children, 71*(2), 149-164. http://dx.doi.org/10.1177/001440290507100202

Gibbs, G. (2006). How assessment frames student learning. In C. Bryan & K. Clegg (Eds.), *Innovative assessment in higher education* (pp. 23–36). Routledge.

Gillham, B. (2008). *Developing a questionnaire*. A&C Black.

Gipps, C. V. (1994). *Beyond testing: Towards a theory of educational assessment.* Psychology Press.

Gipps, C. V. (1999). Socio-Cultural Aspects of Assessment. *Review of Research in Education*, *24*(1), 355–392. https://doi.org/10.3102/0091732x024001355

Gipps, C., & Stobart, C.V. (1997). *Assessment: A teachers' guide to the issues* (3rd ed.). Hodder & Stoughton.

Glaser, R. & Bassok, M . (1989). Learning Theory and the Study of Instruction. *Annual Review of Psychology, 40*, 631-666.

Goldstein, L. M. (2005). *Teacher written commentary: In second language writing classrooms.* University of Michigan Press.

Good, R. (2011). Formative use of assessment information: It's a process, so let's say what we mean. *Practical Assessment, Research, and Evaluation, 16*(1), 3. https://scholarworks.umass.edu/pare/vol16/iss1/3

Gottman, J. M., McFall, R. M., & Barnett, J. T. (1969). Design and analysis of research using time series. *Psychological Bulletin, 72*(4), 299.

Grabe, W., & Kaplan, R. B. (1996). Theory and practice of writing: An applied linguistic perspective. Teaching writing at beginning levels. Longman.

Gray, J. H., & Densten, I. L. (1998). Integrating quantitative and qualitative analysis using latent and manifest variables. *Quality and Quantity, 32*(4), 419-431. http://dx.doi.org/10.1023/a:1004357719066

Greene, J. C., & Caracelli, V. J. (1997). *Advances in mixed-method evaluation: The challenges and benefits of integrating diverse paradigms.* New Directions for Evaluation-Jossey-Bass.

Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis, 11*(3), 255-274. http://dx.doi.org/10.3102/01623737011003255

Greenland, S., Senn, S. J., Rothman, K. J., Carlin, J. B., Poole, C., Goodman, S. N., & Altman, D. G. (2016). Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *European Journal of Epidemiology, 31*(4), 337-350 http://dx.doi.org/10.1007/s10654-016-0149-3

Gregory, K., Cameron, C., & Davies, A. (2000). *Knowing what counts: Self-*

assessment and goal setting. Connections Publishing.

Griffee, D. T. (2018). *An introduction to second language research methods* (2nd ed.). TESL-EJ Publications.

Griffith, A. (2007). *SPSS for Dummies*. John Wiley & Sons.

Guardado, M., & Shi, L. (2007). ESL students' experiences of online peer feedback. *Computers and Composition, 24*(4), 443-461. http://dx.doi.org/10.1016/j.compcom.2007.03.002

Guskey, T. R. (2005). Formative classroom assessment and Benjamin Bloom: Research, theory, and implications', The annual meeting of the American Educational Research Association Montreal, Canada, 11-15. https://files.eric.ed.gov/fulltext/ED490412.pdf

Hafner, J., & Hafner, P. (2003). Quantitative analysis of the rubric as an assessment tool: an empirical study of student peer group rating. *International Journal of Science Education, 25*(12), 1509–1528. http://dx.doi.org/10.1080/0950069022000038268

Halliday, M. (1978b). *Language as social semiotic: The social interpretation of language and meaning.* Arnold.

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English.* Longman English Language Series

Halliday, M., & Hasan, R. (1989). *Language, context, and text: Aspects of language in a social-semiotic perspective* (2nd ed.). Oxford University Press.

Halliday, M.A.K. & Hasan, R. (1976). *Cohesion in English*. London: Longman.

Hamilton, M. L., & Pinnegar, S. (1998). *Reconceptualizing teaching practice: Self-study in teacher education.* Psychology Press.

Hammersley, M. (2007). The issue of quality in qualitative research. *International Journal of Research & Method in Education, 30*(3), 287-305. http://dx.doi.org/10.1080/17437270701614782

Hamp-Lyons, L. (1996). Applying ethical standards to portfolio assessment of writing in English as a second language. In Milanovic, M., & Saville, N. (Eds.) (1996), *Performance testing, cognition and assessment*. (pp. 151-162). Cambridge University Press

Hamp-Lyons, L., & Con, W. (2000). *Assessing the portfolio principles for practice, theory and research.* Hampton Press.

Hamp-Lyons, L., & Reed, R. (1990). *Development of the new Michigan Writing Assessment: Report to the College of LS&A, Mimeo*. English Composition Board.

Hamzaoui-Elachachi, H. (2006). *An exploration into the strategies used for essay writing across three languages: The case of EFL university students* [Doctoral thesis]. University of Tlemcen, Algeria.

Hamzaoui-Elachachi, H. (2010). Development of a writing curriculum for academic purposes at tertiary level: The case of Algerian EFL university students. *ESP World*, *1*(27), 1–9.

Hanrahan, S. J., & Isaacs, G. (2001). Assessing self-and peer-assessment: The students' views. *Higher Education Research & Development*, *20*(1), 53-70. http://dx.doi.org/10.1080/07924360120043658

Harlen, W (2005). Teachers' summative practices and assessment for learning – tensions and synergies. *Curriculum Journal*, *16*(2), 207 223. http://dx.doi.org/10.1080/09585170500136093

Harmer, J. (2004). *How to teach writing*. Pearson Education Limited.

Harrison, K., Joe, O. H., & McNamara, G. (2015). Re-thinking assessment: self-and peer-assessment as drivers of self-direction in learning. *Eurasian Journal of Educational Research*, *15*(60), 75-88.

Hart, A.W. & Reinking, J.A. (1990). *Writing for College and Career* (4th ed.). St. Martin's Press.

Hart, D. (1999). Opening assessment to our students. *Social Education*, *63*(6), 343-45.

Hasan, R. (1996). *Ways of saying: Ways of meaning*. Cassell.

Hassan, B. (2001). The relationship of writing apprehension and self-esteem to the writing quality and quantity of EFL university graduates. *College Education*, 1-37.

Hasyim, S. (2002). Error analysis in the teaching of English. *Ta Lama*, *4*(1), 62-74. kata.petra.ac.id/index.php/ing/article/download/15485/15477

Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. Routledge.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81-112. http://dx.doi.org/10.3102/003465430298487

Hedge, T. (1988). *Writing*. Oxford University Press.

Hendry, G. D., Bromberger, N., & Armstrong, S. (2011). Constructive guidance and feedback for learning: The usefulness of exemplars, marking sheets and different types of feedback in a first year law subject. *Assessment & Evaluation in Higher Education*, *36*(1), 1-11. http://dx.doi.org/10.1080/02602930903128904

Hibbard, K. M., & Wagner, E. (2013). *Assessing and teaching reading composition and writing*, 3-5, (Vol. 4). Routledge.

Hinett, K., and J. Thomas, (1999). *Staff guide to self and peer assessment.* Oxford Centre for Staff and Learning Development

Hinkel, E. (2002). Teaching grammar in writing classes: Tenses and cohesion. In E. Hinkel & S. Fotos (Eds.), *New perspectives on grammar teaching in second language classrooms.* Lawrence Erlbaum.

Hinton, P., Brownlow, C, McMurray, I., & Cozens, B. (2004). *SPSS explained.*

Routledge.

Hopfenbeck, T. N., Flórez Petour, M. T., & Tolo, A. (2015). Balancing tensions in educational policy reforms: Large-scale implementation of assessment for learning in Norway. *Assessment in Education: Principles, Policy & Practice, 22*(1), 44-60. http://dx.doi.org/10.1080/0969594X.2014.996524

Horowitz, D. (1986b). Process, not product: Less than meets the eye. *TESOL quarterly, 20*(1), 141-144. http://dx.doi.org/10.2307/3586397

Horowitz, D. M. (1986). What professors actually require: Academic tasks for the ESL classroom. *TESOL Quarterly, 20*(3), 445-462.

Huberman, A. M., & Miles, M. B. (1994). *Data management and analysis methods.*

Hyland, K. (2003). *Second Language Writing*. Cambridge University Press.

Hyland, F. (2003). Focusing on form: Student engagement with teacher feedback. *System, 31*(2), 217-230. http://dx.doi.org/10.1016/S0346-251X(03)00021-6

Hyland, F., & Hyland, K. (2006). *Feedback in the second language writing*. Cambridge: Cambridge University Press.

Hyland, K. (2007). Genre pedagogy: Language, literacy and L2 writing instruction. *Journal of Second Language Writing, 16*(3), 148-164. http://dx.doi.org/10.1016/j.jslw.2007.07.005

Hyland, K. (2009). *Teaching and researching writing* (2nd ed.). Harlow, Longman

Hyon, S. (1996). Genre in three traditions: Implications for ESL. *TESOL Quarterly, 30*(4), 693-722. http://dx.doi.org/10.2307/3587930

Ibnian, S. S. K. (2017). Writing difficulties encountered by Jordanian EFL learners. *Asian Journal of Humanities and Social Studies, 5*(3).

IELTS, T. (2). Writing band descriptors (public version). Online.[cit. 19-12-13]. Dostupné na: https://takeielts.britishcouncil.org/sites/default/files/IELTS_task_2_Writing_band_descriptors.pdf

Ioannidis, J. P. A. (2005). Why most published research findings are false. PLoS Medicine, 2, *e124*. http://www.plosmedicine.org/article/info:doi/10.1371/journal.pmed.0020124

Jackson, C. W., & Larkin, M. J. (2002). Teaching students to use grading rubrics. *Teaching Exceptional Children, 35*(1), 40-45. http://dx.doi.org/10.1177/004005990203500106

Jacobs, H. L., Kahn, H. D., Stralka, K. A., & Phan, D. B. (1998). Estimates of per capita fish consumption in the US based on the continuing survey of food intake by individuals (CSFII). *Risk Analysis, 18*(3), 283-291. http://dx.doi.org/10.1111/j.1539-6924.1998.tb01295.x

James, C. (2013). *Errors in language learning and use: Exploring error analysis*. Routledge.

James, M. (1992). Assessment for learning. Annual Conference of The Association For The Supervision And Curriculum Development (Assembly Session On Critique Of Reforms In Assessment and Testing in Britain). New Orleans, LA.

James, M. (1998). *Using assessment for school improvement*. Heinemann Educational.

Jessup, L. M., Connolly, T., & Tansik, D. A. (1990). Toward a theory of automated group work: The deindividuating effects of anonymity. *Small Group Research, 21*(3), 333-348. http://dx.doi.org/10.1177/1046496490213003

Jick, T. D. (1979). Mixing qualitative and quantitative methods: Triangulation in action. *Administrative Science Quarterly, 24*(4), 602–611. http://dx.doi.org/10.2307/2392366

Joffe, H. (2012). Thematic analysis. In Harper, D., & Thompson, A. R. (Eds.) *Qualitative research methods in mental health and psychotherapy: A guide for students and practitioners*, (pp, 210-223). John Wiley & Sons.

Joffe, H., & Yardley, L. (2004). Content and thematic analysis. Research methods for clinical and health psychology. In Marks, D. F., & Yardley, L. (Eds.). (2004). *Research methods for clinical and health psychology* (pp. 65-68). Sage Publications.

Jönsson, A., & Panadero, E. (2017). The use and design of rubrics to support assessment for learning. In Carless, D., Bridges, S. M., Chan, C. K. Y., & Glofcheski, R. (Eds.), *Scaling up assessment for learning in higher education*, (pp. 99-111). Singapore.

Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review, 2*(2), 130-144. http://dx.doi.org/10.1016/j.edurev.2007.05.002

Joseph, S., Rickett, C., Northcote, M., & Christian, B. J. (2020). Who are you to judge my writing? : Student collaboration in the co-construction of assessment rubrics. *New Writing, 17*(1), 31-49. https://doi.org/10.1080/14790726.2019.1566368

Kaplan, R. B. (1966). Cultural thought patterns in inter-cultural education. *Language Learning*, *16*(1-2), 1–20. http://dx.doi.org/10.1111/j.1467-1770.1966.tb00804.x

Kathpalia, S. S., & Heah, C. (2010). Sharing The Responsibility Of Feedback In Academic Writing. *English Teacher*, 39.

Kaufman, J. K., & Schunn, C. D. (2011). Students' perceptions about peer assessment for writing: Their origin and impact on revision work. *Instructional Science, 39*(3), 387–406. http://dx.doi.org/10.1007/s11251-010-9133-6

Kelle, U., & Erzberger, C. (2003). *Making inferences in mixed methods: The rules of integration.* Thousand Oaks

Kemper, D., Meyer, V., Van Rys, J., & Sebranek, P. (2018). *Fusion: Integrated reading and writing*. Cengage Learning.

Kepner, C., (1991). An experiment in the relationship of types of written feedback to the development of writing skills. *Modern Language Journal 75*(3), 305–313. http://dx.doi.org/10.2307/328724

Kepner, C.G. (1991). An Experiment in the Relationship of Types of Written Feedback to the Development of Second Language Writing Skills. *The Modern Language Journal*, *75*(3), 305-313. http://www.jstor.org/stable/328724

Kerssen-Griep, J., & Witt, P. (2012). Instructional feedback II: How do instructor immediacy cues and facework tactics interact to predict student motivation and fairness perceptions? *Communication Studies*, *63*(4), 498-517. http://dx.doi.org/10.1080/10510974.2011.632660

Khanchali, M. (2017). *Collaborative teaching as a treatment for the low proficiency of students in writing: The case of second year students at Batna University* [Doctoral Thesis]. University of Batna 2.

Khouya, Y. B. (2018). Students Demotivating Factors in the EFL Classroom: The Case of Morocco. *Advances in Language and Literary Studies*, *9*(2), 150-159.

Kift, S., & Nelson, K., (2005). Beyond curriculum reform: Embedding the transition experience. Conference Proceedings.

Kirmizi, Ö., & Aydin, E. (2019). A qualitative analysis of writing problems of low achieving advanced learners in Turkey. *Current Research in Education*, *5*(1), 34-54.

Klenowski, V. (1995). Student Self-Evaluation Processes in Student-Centred Teaching and Learning Contexts of Australia and England. *Assessment in Education: Principles, Policy & Practice*, *2*(2), 145–163. https://doi.org/10.1080/0969594950020203

Klenowski, V. (2002). *Developing portfolios for learning and assessment: Processes and principles*. Routledge.

Klenowski, V. (2009). Assessment for learning revisited: An Asia-Pacific perspective. *Assessment in Education: Principles, Policy & Practice*, *16* (3), 263-268. http://dx.doi.org/10.1080/09695940903319646

Klenowski, V., Askew S., & Carnell, E. (2006). Portfolios for learning, assessment and professional development in higher education. *Assessment and Evaluation in Higher Education*, *31*(3), 267-286. http://dx.doi.org/10.1080/02602930500352816

Knapp, H. (2017). *Introductory statistics using SPSS* (2nd ed.). Sage Publications.

Knapp, P., & Watkins, M. (1994). *Context, text, grammar*. Broadway: Text Productions.

Knight, P. T. (2002). The achilles' heel of quality: The assessment of student learning.

Quality in Higher Education, 8(1). 107-115. https://doi.org/10.1080/13538320220127506

Krause, K., Hartley, R., James, R., McInnis, C. (2005). The first year experience in Australian universities: Findings from a decade of national studies. Centre for the Study of Higher Education, University of Melbourne.

Kroll, B. (1990). *Second language writing*. Cambridge University Press.

Kulhavy, R. W. (1977). Feedback in Written Instruction. *Review of Educational Research, 47*(2), 211–232. http://dx.doi.org/10.3102/00346543047002211

Kulhavy, R. W., & Wager, W. (1993). Feedback in programmed instruction: Historical context and implications for practice. In J. V. Dempsey, & G. C. Sales (Eds.), *Interactive instruction and feedback* (pp. 3–20). Englewood Cliffs, NJ: Educational Technology.

Kumar, R. (2018). *Research methodology: A step-by-step guide for beginners* (3rd ed.). Sage publications.

Kumaravadivelu, B. (2003). *Beyond methods: Macrostrategies for language teaching*. Yale University Press.

Kumaravadivelu, B. (2003). A postmethod perspective on English language teaching. *World Englishes, 22*(4), 539-550.

Kusuran, A. (2017). L2 English spelling error analysis: An investigation of English spelling errors made by Swedish senior high school students. Humanities and Social Sciences, diva portal.org. http://www.divaportal.org/smash/get/diva2:1078118/FULLTEXT01.pdf

Lalande, J.F., (1982). Reducing composition errors. An experiment. *Modern Language Journal 66*(2), 140–149. http://www.jstor.org/stable/326382

Lam, R. (2015). Convergence and divergence of process and portfolio approaches to L2 writing instruction: Issues and implications. *RELC Journal, 46*(3), 293-308. http://dx.doi.org/10.1177/0033688215597119

Lam, R. (2018). Promoting self-reflection in writing: A showcase portfolio approach. In Burns, A., & Siegel, J. (Eds.), *International perspectives on teaching the four skills in ELT* (pp. 219-231). Palgrave Macmillan, Cham.

Lam, R. (2018). *Portfolio assessment for the teaching and learning of writing*. Springer.

Lam, R., & Lee, I. (2009). Balancing the dual functions of portfolio assessment. *ELT Journal, 64*(1), 54-64. http://dx.doi.org/10.1093/elt/ccp024

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation* (18th ed.). Cambridge University Press.

Laveault, D., & Allal, L. (2016). Implementing assessment for learning: Theoretical and practical issues. In Laveault, D., & Allal, L. (Eds.), *Assessment for learning: Meeting the challenge of implementation* (pp.1-18). Springer.

Leach, L. (2012). Optional self-assessment: some tensions and dilemmas. *Assessment & Evaluation in Higher Education, 37*(2), 137-147. http://dx.doi.org/10.1080/02602938.2010.515013

Lee, G., & Schallert, D. L. (2008). Constructing trust between teacher and students through feedback and revision cycles in an EFL writing classroom. *Written Communication, 25*(4), 506-537. http://dx.doi.org/10.1177/0741088308322301

Lee, I. (2002). Teaching coherence to ESL students: a classroom inquiry. *Journal of Second Language Writing, 11*(2), 135-159. http://dx.doi.org/10.1016/s1060-3743(02)00065-6

Lee, I. (2017). *Classroom writing assessment and feedback in L2 school contexts*. Springer Singapore.

Lewin, K. (1952). Group decision and social change. In G. E. Swanson,T. M. Newcomb, & F. E. Hartley (Eds.), *Readings in social psychology*. Holt.

Li, H., Xiong, Y., Hunter, C. V., Guo, X., & Tywoniw, R. (2019). Does peer assessment promote student learning? A meta-analysis. *Assessment & Evaluation in Higher Education, 45*(2), 193-211. http://dx.doi.org/10.1080/02602938.2019.1620679

Li, H., Xiong, Y., Zang, X., L. Kornhaber, M., Lyu, Y., Chung, K. S., & K. Suen, H. (2015). Peer assessment in the digital age: A meta-analysis comparing peer and teacher ratings. *Assessment & Evaluation in Higher Education, 41*(2), 245-264. http://dx.doi.org/10.1080/02602938.2014.999746

Li, L. (2017). The role of anonymity in peer assessment. *Assessment and Evaluation in Higher Education, 42*(4), 645-656. http://dx.doi.org/10.1080/02602938.2016.1174766

Li, L., & Gao, F. (2016). The effect of peer assessment on project performance of students at different learning levels. *Assessment & Evaluation in Higher Education, 41*(6), 885-900. http://dx.doi.org/10.1080/02602938.2015.1048185

Li, L., & Grion, V. (2019). The power of giving feedback and receiving feedback in peer assessment. *All Ireland Journal of Higher Education, 11*(2).

Li, L., Liu, X., & Steckelberg, A. L. (2010). Assessor or assessee: How student learning improves by giving and receiving peer feedback. *British Journal of Educational Technology, 41*(3), 525-536. http://dx.doi.org/10.1111/j.1467-8535.2009.00968.x

Li, L., Liu, X., & Zhou, Y. (2012). Give and take: A re-analysis of assessor and assessee's roles in technology-facilitated peer assessment. *British Journal of Educational Technology, 43*(3), 376-384. http://dx.doi.org/10.1111/j.1467-8535.2011.01180.x

Lincoln, Y. S., Guba, G. E. (1985). *Naturalistic inquiry*. London Sage Publications.

Liu, M., & Braine, G. (2005). Cohesive features in argumentative writing produced by

Chinese undergraduates. *System, 33*(4), 623-636. http://dx.doi.org/10.1016/j.system.2005.02.002

Logan, E. (2009). Self and Peer Assessment in Action. *Practitioner Research in Higher Education, 3*(1), 29-35.

Looney, J. (2008). *Teaching, learning and assessment for adults*. Improving foundation.

Lu, R., & Bol, L. (2007). A comparison of anonymous versus identifiable e-peer review on college student writing performance and the extent of critical feedback. *Journal of Interactive Online Learning, 6*(2). https://digitalcommons.odu.edu/efl_fac_pubs/5

Lundstrom, K., & Baker, W. (2009). To give is better than to receive: The benefits of peer review to the reviewer's own writing. *Journal of Second Language Writing, 18*(1), 30-43. http://dx.doi.org/10.1016/j.jslw.2008.06.002

Mackey, A., & Gass, S. M. (2005). *Second Language Research: Methodology and design*. Routledge.

MacLeod, L. (1999). Computer-aided peer review of writing. *Business Communication Quarterly, 62*(3), 87-95. http://dx.doi.org/10.1177/108056999906200309

Mangelsdorf, K. (1992). Peer reviews in the ESL composition classroom: What do the students think?. *ELT journal*, *46*(3), 274-284.

Martin, J. R. (1984). Language, register and genre. In F. Christie (Ed.), *Children Writing: Reader* (pp. 21-29). Deakin University Press.

Marx, G. T. (1999). What's in a name? Some reflections on the sociology of anonymity. *Information Society, 15*(2), 99–112. http://dx.doi.org/10.1080/019722499128565

Matsuda, M. (1999). Interlanguage pragmatics: What can it offer to language teachers. *The CATESOL Journal, 11*(1), 39-59.

Matsuda, P. K. (2003). Second language writing in the twentieth century: A situated historical perspective. In B. Kroll (Ed.), *Exploring the dynamics of second language writing* (pp. 15-34). Cambridge University Press.

Matsuda, P. K., Ortmeier-Hooper, C., & Matsuda, A. (2009). *The expansion of second language writing.* Sage handbook of writing development.

Matsumoto, K. (1995). Research paper writing strategies of professional Japanese EFL writers. *TESL Canada Journal, 13*(1), 17–27. http://dx.doi.org/10.18806/tesl.v13i1.658

Maxwell, J. A., & Mittapalli, K. (2010). Realism as a stance for mixed methods research. In Tashakkori, A., & Teddlie, C. (Ed.), *Sage handbook of mixed methods in social & behavioral research* (pp. 145-168). SAGE publications

McCarter, S., & Whitby, N. (2014). *Improve your IELTS writing skills.* Macmillan

Publishers Limited.

McCleary, R., McDowall, D., & Bartos, B. J. (2017). *Design and analysis of time series experiments*. Oxford University Press.

McComiskey, B. (2000). *Teaching composition as a social process*. Logan, UT: Utah State University Press.

McConnell, K. D. (2013). Rubrics as catalysts for collaboration: a modest proposal. *European Journal of Higher Education, 3*(1), 74-88. http://dx.doi.org/10.1080/21568235.2013.778043

McFarland, E. (2015). Importance of handwriting: how Montessori didactic materials support handwriting. [Doctoral Dissertation]. The University of Wisconsin - River Falls

McInnes, C., & James, R. (1995). *First year on campus: Diversity in the initial experiences of Australian undergraduates.* AGPS, Canberra.

McNamara, M. J., & Deane, D. (1995). Self-Assessment Activities: Toward Language Autonomy in Language Learning. *TESOL Journal, 5*(1), 17-21.

Messick, S. (1989). Validity. In R. L. Linquist (Ed.), *Educational measurement* (3rd ed.), (pp.13-103). Macmillan.

Meyer, L. (1986). Strategies for Correcting Students' Wrong Responses. *The Elementary School Journal*, *87*(2), 227–241. http://dx.doi.org/10.1086/461491

Miller, A. H., Imrie, B. W., & Cox, K. (1998). *Student Assessment in Higher Education*. Kogan Page.

Miller, C. R. (1984). Genre as social action. *Letras & Letras, 31*(3), 56-72.

Mohammad, T., & Hazarika, Z. (2016). Difficulties of learning EFL in KSA: Writing skills in context. *International Journal of English Linguistics, 6*(3), 105-117.

Montaner-Villalba, S. (2019). Critical Thinking and Written Production in the EFL classroom through Blogging. International Insights: Equality in Education, 167.

Montgomery, K. (2000). Classroom rubrics: Systematizing what teachers do naturally. *The clearing house*, *73*(6), 324-328. https://www.jstor.org/stable/30189610

Montgomery, K. (2002). Authentic Tasks and Rubrics: Going beyond Traditional Assessments in College Teaching. *College Teaching 50* (1), 34–40. http://dx.doi.org/10.1080/87567550209595870

Morse JM (2008). Confusing Categories and Themes. *Qualitative Health Research, 18*(6), 727-728. http://dx.doi.org/10.1177/1049732308314930

Moskal, B. M., & Leydens, J. A. (2000). Scoring rubric development: Validity and reliability. *Practical assessment, research, and evaluation, 7*(1), 10.

Moss, P. A. (1994). Can there be validity without reliability?. *Educational Researcher*, *23*(2), 5-12. http://dx.doi.org/10.3102/0013189X023002005

Moussaoui, S. (2012). An investigation of the effects of peer evaluation in enhancing Algerian students' writing autonomy and positive affect. *Procedia—Social and Behavioral Sciences*, *69*, 1775–1784.

Mowl, G., & Pain, R. (1995). Using self and peer assessment to improve students' essay writing: A case study from geography. *Innovations in Education and Training International*, *32*(4), 324-335. http://dx.doi.org/10.1080/1355800950320404

Murphy, R. (2006). Evaluating new priorities for assessment in higher education. In Bryan, C. & Clegg, K. (Eds.), *Innovative Assessment in Higher Education* (pp. 37-47). Routledge.

Murray, D. M. (2004). Following the Voice of the Draft. In Behar, R., & Bleich, D. (Eds.), *Autobiographical writing across the disciplines: A reader* (pp, 129-137). Duke University Press.

Myhill, D. (2009). *From talking to writing: Linguistic development in writing*. The British Psychological Society.

Ndoye, A. (2017). Peer/Self Assessment and Student Learning. *International Journal of Teaching and Learning in Higher Education*, *29*(2), 255-269.

Needleman, C., & Needleman, M. L. (1996). Qualitative methods for intervention research. *American Journal of Industrial Medicine*, *29*(4), 329-337. http://dx.doi.org/10.1002/(sici)1097-0274(199604)29:4

Nelms, K. R. (2015). Technology Knowledge Self-Assessment and Pre-Test Performance among Digital Natives. *Journal of Learning in Higher Education*, *11*(1), 99-107.

Nelson, M. M., & Schunn, C. D. (2009). The nature of feedback: How different types of peer feedback affect writing performance. *Instructional Science*, *37*(4), 375-401. http://dx.doi.org/10.1007/s11251-008-9053-x

Nelson, R. H., & Trafford, C. H. (2003). Handwriting a complete guide to instruction. Peterson Directed Handwriting. https://www.petersonhandwriting.com/Publications/PDF_versions/ReviewSections123.pdf

Neuendorf, K. A. (2019). Content analysis and thematic analysis. In Brough, P. (Ed.), *Advanced research methods for applied psychology* (pp. 211-223). Routledge.

Neuman, W. L. (2014). *Social research methods: Qualitative and quantitative approaches* (7th ed.) Pearson Education.

Nicol, D. (2008). *Rethinking assessment and feedback*. Presentation at University of Glasgow.

Nicol, D. (2010). From monologue to dialogue: Improving written feedback processes in mass higher education. *Assessment and Evaluation in Higher Education*, *35*(5), 501–517. http://dx.doi.org/10.1080/02602931003786559

Nicol, D. J., & 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. http://dx.doi.org/10.1080/03075070600572090

Nilson, L. B. (2003). Improving student peer feedback. *College Teaching*, *51*(1), 34-38 http://dx.doi.org/10.1080/87567550309596408

Ninness, H. A., Ellis, J., & Ninness, S. K. (1999). Self-assessment as a learned reinforcer during computer interactive math performance: An experimental analysis. *Behavior Modification*, *23*(3), 403-418. http://dx.doi.org/10.1177/0145445599233004

Nitko, A. J. (2001). Conceptual frameworks to accommodate the validation of rapidly changing requirements for assessments. *Curriculum and Assessment*, *1*, 143-163.

Nitko, A. J., & Brookhart, S. M. (2008). *Assessment and grading in classrooms*. Pearson/Merrill Prentice Hall.

Northwest Regional Educational Laboratory. (2008b). 6+1 Trait® Writing Scoring Guide- 5 Point Condensed.

Nunan, D. (1992). *Research methods in language learning*. Cambridge University Press.

O'Donnell, J. A., Oakley, M., Haney, S., O'Neill, P. N., & Taylor, D. (2011). Rubrics 101: a primer for rubric development in dental education. *Journal of dental education*, *75*(9), 1163-1175. http://dx.doi.org/10.1002/j.0022-0337.2011.75.9.tb05160.x

O'Donnell, T.D. & Paiva, J.L. (1993). *Independent Writing*. (2nd ed.). Heinle & Heinle Publishers.

Office of Institutional Research and Assessment (2017). Using Rubrics To Assess Student Learning Outcomes At The Program Level. https://oira.unc.edu/wp-content/uploads/sites/297/2017/07/Developing-and-Using-Rubrics.pdf

Okilwa, N. S., & Shelby, L. (2010). The effects of peer tutoring on academic performance of students with disabilities in grades 6 through 12: A synthesis of the literature. *Remedial and Special Education*, *31*(6), 450-463. http://dx.doi.org/10.1177/0741932509355991

Orsmond, P, Reiling, K., & Merry, S (2002). The student use of tutor formative feedback in their learning. Paper presented at the Learning Communities and Assessment Cultures Conference organised by the EARLI Special Interest Group on Assessment and Evaluation, University of Northumbria

Oshima, A., & Hogue, A. (2009). *Writing Academic English* (4th ed.). Upper Sadder River: Pearson Educations.

Ourghi, R. (2002). *The implications of critical reflection and discourse awareness for educational change: The case of the writing curriculum, learner, and teacher development at the university level* [Doctoral thesis]. University of Tlemcen, Algeria.

Pajares, F., & Valiante, G. (1999). Grade level and gender differences in the writing self-beliefs of middle school students. *Contemporary Educational Psychology*, *24*(4), 390-405. http://dx.doi.org/10.1006/ceps.1998.0995

Paltridge, B. (2016). Data selection as an ethical issue: Dealing with outliers in telling a research story. In De Costa (Ed.), *Ethics in applied linguistics research* (pp. 38-50). Routledge.

Panadero, E. (2016). Is it safe? Social, interpersonal, and human effects of peer assessment. In Brown, G, & Harris, L., *Handbook of human and social conditions in assessment*, (pp. 247-266). Routledge.

Panadero, E., & Alonso-Tapia, J. (2013). Self-assessment: theoretical and practical connotations, when it happens, how is it acquired and what to do to develop it in our students. *Electronic Jounral of Research in Educational Psychology, 11* (2), 551-576.

Panadero, E., & Alqassab, M. (2019). An empirical review of anonymity effects in peer assessment, peer feedback, peer review, peer evaluation and peer grading. *Assessment & Evaluation in Higher Education.* 1–26. http://dx.doi.org/10.1080/02602938.2019.1600186

Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational research review, 9*, 129-144. http://dx.doi.org/10.1016/j.edurev.2013.01.002

Panadero, E., & Romero, M. (2014). To rubric or not to rubric? The effects of self-assessment on self-regulation, performance and self-efficacy. *Assessment in Education: Principles, Policy & Practice*, *21*(2), 133-148. http://dx.doi.org/10.1080/0969594X.2013.877872

Panadero, E., Alonso-Tapia, J., & Reche, E. (2013). Rubrics vs. self-assessment scripts effect on self-regulation, performance and self-efficacy in pre-service teachers. *Studies in Educational Evaluation*, *39*(3), 125-132.

Panadero, E., Andrade, H., & Brookhart, S. (2018). Fusing self-regulated learning and formative assessment: A roadmap of where we are, how we got here, and where we are going. *Australian Educational Researcher*, *45*(1), 13-31. http://dx.doi.org/10.1007/s13384-018-0258-y

Panadero, E., Brown, G. T. L., & Strijbos, J. W. (2016). The future of student self-assessment: A review of known unknowns and potential directions. *Educational Psychology Review, 28*(4), 803-830. http://dx.doi.org/10.1007/s10648-015-9350-2

Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review*, *22*, 74-98. https://doi.org/10.1016/j.edurev.2017.08.004

Panadero, E., Jonsson, A., & Strijbos, J. W. (2016). Scaffolding self-regulated learning through self-assessment and peer assessment: Guidelines for classroom implementation. In Laveault, D., & Allal, L. (Eds.), *Assessment for learning: Meeting the challenge of implementation* (pp. 311-326). Springer.

Panadero, E., Lipnevich, A., & Broadbent, J. (2019). Turning self-assessment into self-feedback. In Henderson, M., Ajjawi, R., Boud, D., & Molloy, E. (Eds.), *The impact of feedback in higher education* (pp. 147-163). Springer Nature.

Panadero, E., Romero, M., & Strijbos, J. W. (2013). The impact of a rubric and friendship on peer assessment: Effects on construct validity, performance, and perceptions of fairness and comfort. *Studies in Educational Evaluation*, *39*(4), 195-203. http://dx.doi.org/10.1016/j.stueduc.2013.10.005

Park, T. (2003). Scoring procedures for assessing writing. *Studies in Applied Linguistics and TESOL*, *3*(1), 1-3.

Parks, A. F., Levemier, J. A., & Hollowell, I. M. (1986). *Structuring paragraphs: A guide to effective writing* (2nd ed.). St. Martin's Press.

Patton, M. Q. (1990). *Qualitative evaluation and research methods*. SAGE Publications.

Paulson, F. L. (1991). What Makes a Portfolio a Portfolio? *Educational Leadership*, *48*(5), 60-63.

Payne, D. A. (2003). *Applied educational assessment* (2nd ed). Wadsworth/Thomson Learning.

Peirce, B. N., Swain, M., & Hart, D. (1993). Self-assessment, French immersion, and locus of control. *Applied Linguistics*, *14*(1), 25-42. http://dx.doi.org/10.1093/applin/14.1.25

Popham, W. J. (1997). What's Wrong--and What's Right--with Rubrics. *Educational leadership*, *55*(2), 72-75.

Prodromou. L. (1995). The Backwash effect: From testing to teaching. *English Language Journal*, *21*(1), 1-25. http://dx.doi.org/10.1093/elt/49.1.13

Prohaska, V., & Maraj, F. (1995). Low and medium ability students confidently overestimate all their grades. Paper presented at the Seventh. Annual Convention of the American Psychological Society.

Pui, P., Yuen, B., & Goh, H. (2020). Using a criterion-referenced rubric to enhance student learning: a case study in a critical thinking and writing module. *Higher Education Research & Development*, *40*(5), 1056-1069. http://dx.doi.org/10.1080/07294360.2020.1795811

Race, P. (2001). A briefing on self, peer and group assessment. LTSN Generic Centre Assessment Guides Series.

Raimes, A. (1983). *Techniques in teaching writing.* Oxford University Press.

Raimes, A. (1985). What unskilled ESL students do as they write: A classroom study of composing. *TESOL Quarterly, 19*(2), 229-258.

Raimes, A. (1987). Language proficiency, writing ability, and composing strategies: A study of ESL college student writers. *Language learning, 37*(3), 439-468. http://dx.doi.org/10.1111/j.1467-1770.1987.tb00579.x

Raimes, A. (1993). Out of the woods: Emerging traditions in the teaching of writing. In S. Silberstein (Ed.), *State o f the art TESOL essays: 25 years o f the discipline* (pp. 237-260). TESOL Inc.

Rakedzon, T., & Baram-Tsabari, A. (2017). To make a long story short: A rubric for assessing graduate students' academic and popular science writing skills. *Assessing Writing, 32*, 28-42. http://dx.doi.org/10.1016/j.asw.2016.12.004

Rakrak, M. (2020). EFL Writing Instruction Research in Morocco: A Historical Overview. *Journal of World Englishes and Educational Practices, 2*(6), 23-28.

Ramaprasad, A. (1983). On the definition of feedback. *Behavioral Science, 28*(1), 4-13. http://dx.doi.org/10.1002/bs.3830280103

Rass, R. A. (2015). Challenges face Arab students in writing well-developed paragraphs in English. *English Language Teaching, 8*(10), 49-59.

Reddy, Y. M., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment & Evaluation in Higher Education, 35*(4), 435-448. http://dx.doi.org/10.1080/02602930902862859

Reid, J. (1993). *Teaching ESL writing*. Englewood Cliffs, NJ: Regents/ Prentice Hall.

Reid, J. (1994). Responding to ESL students' texts: The myths of appropriation. *TESOL Quarterly, 28*(2), 272-292. http://www.jstor.org/stable/3587434

Reynolds-Keefer, L. (2010). Rubric-referenced assessment in teacher preparation: An opportunity to learn by using. *Practical Assessment, Research, and Evaluation, 15*(1), 8.

Rezaei, A. R., & Lovorn, M. (2010). Reliability and validity of rubrics for assessment through writing. *Assessing writing, 15*(1), 18-39.

Robinson, R. S. (2014). Purposive sampling. In Michalos, A C (Eds.), *Encyclopedia of quality of life and well-being research* (pp. 233-249).Springer.

Rolheiser, C., Bower, B., & Stevahn, L. (2000). *The portfolio organizer: Succeeding with portfolios in your classroom.* ASCD.

Rose, D., & Martin, J. R. (2012). *Learning to write, reading to learn: Genre, knowledge and pedagogy in the Sydney School.* Equinox.

Rotsaert, T., Panadero, E., & Schellens, T. (2018). Anonymity as an instructional scaffold in peer assessment: its effects on peer feedback quality and evolution in students' perceptions about peer assessment skills. *European Journal of Psychology of Education, 33*(1), 75-99. http://dx.doi.org/10.1007/s10212-017-0339-8

Rowntree, D. (1987). *Assessing Students: How shall we know them?* (2nd ed.). Kogan Page.

Ruegg, R., & Sugiyama, Y. (2013). Organization of ideas in writing: what are raters sensitive to?. *Language Testing in Asia, 3*(1), 1-13. http://dx.doi.org/10.1186/2229-0443-3-8

Russell, D., Lea, M., Parker, J., Street, B., & Donahue (2009). Exploring notions of genre in 'academic literacies' and 'writing across the curriculum': Approaches across countries and contexts. In C. Bazerman, A. Bonini & D. Figueiredo (Eds.), *Genre in a changing world: Perspectives on writing* (pp. 459-491). WAC Clearinghouse/Parlor Press.

Russell, T. A., Qualter, A. & Mcguigan, L. (1995). Reflections on the implementation of National evaluation study in England. *International Journal of Science Education, 17*(4), 481–492. https://doi.org/10.1080/0950069950170407

Sadler, D. R. (2009). Indeterminacy in the use of preset criteria for assessment and grading. *Assessment & Evaluation in Higher Education, 34*(2), 159-179. http://dx.doi.org/10.1080/02602930801956059

Sadler, D. R. (2010). Beyond feedback: Developing student capability in complex appraisal. *Assessment & Evaluation in Higher Education, 35*(5), 535-550. http://dx.doi.org/10.1080/02602930903541015

Salim, S., & Suppramaniam, J. M. (2020). Validating Student Voice in Writing through Self-Assessment and Rubrics Creation. In Tan, K. (Ed.), *Introduction to decoding rubrics*. (pp. 109-121). Routledge.

Salkind, N. J. (2018). *Exploring research*. Pearson Educación.

Sambell, K., McDowell, L., & Montgomery, C. (2013). Assessment *for learning in higher education*. Routledge.

Sanchez, C. E., Atkinson, K. M., Koenka, A. C., Moshontz, H., & Cooper, H. (2017). Self-grading and peer-grading for formative and summative assessments in 3rd through 12th grade classrooms: A meta-analysis. *Journal of Educational Psychology, 109*(8), 1049.

Sandelowski, M., & Barroso, J. (2003). Classifying the findings in qualitative studies. *Qualitative Health Research, 13*(7), 905-923. http://dx.doi.org/10.1177/1049732303253488

Sandelowski, M., & Barroso, J. (2007). Handbook for synthesizing qualitative research. Springer Publishing Company.

Schon, D. A. (1983). The Reflective Practitioner. Basic Books.

Scriven, M. (1967). The methodology of evaluation. *AERA Monograph Series on Curriculum Evaluation, 1*, 39-83.

Scriven, M. (1991). Beyond Formative and Summative Evaluation., In Mclaughlin, M. & Phillips, D. C. (Eds.), *Evaluation and education: At quarter century*. The

National Society for the Study of Education.

Scriven, M. (1996). Types of Evaluation and Types of Evaluator. *Evaluation Practice. 17*(2), 151-161.

Sebba, J., Deakin Crick, R., Yu, G., Lawson, H. & Harlen, W. (2008). Systematic Review of Research Evidence of the Impact on Students in Secondary Schools of Self and Peer-assessment Report.

Seliger, H. W, & Shohamy, E. (1989). *Second language research methods*. Oxford University Press.

Senge, P. (2000). *Schools that learn*. Doubleday.

Seow, A. (2002). The writing process and process writing. In Jack Croft Richards, J, C., & Renandya, W., A., *Methodology in language teaching: An anthology of current practice* (pp. 315-320). Cambridge University Press.

Shadish, W. R., T. D. Cook, D. T. Campbell. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Houghton Mifflin.

Shavelson, R. J., Black, P., Wiliam, D., & Coffey, J. (2007). On linking formative and summative functions in the design of large-scale assessment systems. Educational Evaluation and Policy Analysis. http://www.stanford.edu/dept/SUSE/SEAL/Reports_Papers/On%20Aligning%20Formative%20and%20Summative%20Functions_Submit.doc

Shen, B., Bai, B., & Xue, W. (2019). The effects of peer assessment on learner autonomy: An empirical study in a Chinese college English writing class. Studies in *Educational Evaluation*, *64*, 100821. http://dx.doi.org/10.1016/j.stueduc.2019.100821

Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for Information*, *22*(2), 63-75. http://dx.doi.org/10.3233/EFI-2004-22201

Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, *29*(7), 4-14. https://doi.org/10.3102/0013189X029007004

Shokrpour, N., & Fallahzadeh, M. H. (2007). A survey of the students and interns' EFL writing problems in Shiraz University of Medical Sciences. *Asian EFL Journal*, *9*(1), 147-163.

Shvidko, E. (2020). Taking into account interpersonal aspects of teacher feedback: Principles of responding to student writing. Journal *on Empowering Teaching Excellence*, *4*(2), 7. https://digitalcommons.usu.edu/jete/vol4/iss2/7

Silva, T. (1990). Second language composition instruction: Developments, issues, and directions in ESL. In. B. Kroll (Ed.), *Second language writing: Research insights for the classroom* (pp. 11-23). Cambridge University Press.

Simpson, J. M. (2000). Topical structure analysis of academic paragraphs in English and Spanish. *Journal of Second Language Writing*, *9*(3), 293-309.

http://dx.doi.org/10.1016/s1060-3743(00)00029-1

Sinclair, R.S. (2010). Thinking and writing: Cognitive science and intelligence analysis. Center for the Study of Intelligence of Washington, DC.

Smith, F. (2013). *Writing and the writer*. Routledge.

Smith, K., & Tillema, H. (2003). Clarifying different types of portfolio use. *Assessment & Evaluation in Higher Education*, *28*(6), 625-648. http://dx.doi.org/10.1080/0260293032000130252

Sommers, N. (2013). *Responding to student writers*. Bedford/St. Martin's.

Spandel, V. (2006). In defense of rubrics. *English Journal*, *96*(1), 19.

Spendlove, D. (2009). *Putting assessment for learning into practice.* A&C Black.

Spiller, D. (2012). Assessment matters: Self-assessment and peer assessment. The *University of Waikato*, *13*, 2-18.

Stallings, V., & Tascione, C. (1996). Student self-assessment and self-evaluation. *The Mathematics Teacher*, *89*(7), 548-554.

Sternberg, R. J. (1998). Metacognition, abilities, and developing expertise: What makes an expert student? *Instructional Science*, *26*(1), 127-140. https://doi.org/10.1023/a:1003096215103

Stevens, D. D., & Levi, A. J. (2013). *Introduction to rubrics: An assessment tool to save grading time, convey effective feedback, and promote student learning*. Stylus Publishing, LLC.

Stiggins, R. (2007). Assessment through the student's eyes. *Educational leadership*, *64*(8), 22.

Stiggins, R. J. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan*, *83*(10), 758-765. https://doi.org/10.1177/003172170208301010

Stiggins, R. J. (2009). Essential FA competencies for teachers and school leaders. In H. L. Andrade & G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 233–250). Routledge.

Stiggins, R., & Chappuis, J. (2006). What a difference a word makes. *Journal of Staff Development*, *27*(1), 10-14.

Stone, D. H. (1993). Design a questionnaire. British Medical Journal, 307(6914), 1264-1266. http://dx.doi.org/10.1136/bmj.307.6914.1264

Storch, N. (2009). The impact of studying in a second language (L2) medium university on the development o f L2 writing. *Journal of Second Language Writing*, *18*(2), 103-118. http://dx.doi.org/10.1016/j.jslw.2009.02.003

Strijbos, J. W., & Wichmann, A. (2018). Promoting learning by leveraging the collaborative nature of formative peer assessment with instructional scaffolds. *European Journal of Psychology of Education*, *33*(1), 1-9. http://dx.doi.org/10.1007/s10212-017-0353-x

Strijbos, J. W., Ochoa, T. A., Sluijsmans, D. M. A., Segers, M., & Tillema, H. H. (2009). Fostering interactivity through formative peer assessment in (web-based) collaborative learning environments. In C. Mourlas, N. Tsianos, & P. Germanakos (Eds.), *Cognitive and emotional processes in Web-Based education: Integrating human factors and personalization* (pp. 375-395). IGI Global

Sullivan, M. (2018). *Statistics: Informed decisions using data* (.5th ed.) Pearson.

Sundeen, T. H. (2014). Instructional rubrics: Effects of presentation options on writing quality. *Assessing writing, 21*, 74-88. http://dx.doi.org/10.1016/j.asw.2014.03.003

Swales, J. (1990). *Genre analysis: English in academic and research settings*. Cambridge University Press.

Taras, M. (1999). Student self-assessment as a means of promoting students autonomy and independence. In M. Taras (Ed.), *Innovations in learning and teaching: Teaching Fellowships at the University of Sunderland* (pp. 61–83). University of Sunderland Press

Taras, M. (2005). Assessment - Summative and Formative - Some theoretical reflections. *British Journal of Educational Studies, 53*(4), 466-478. https://doi.org/10.1111/j.1467-8527.2005.00307.x

Taras, M. (2007a). Machinations of assessment: metaphors, myths and realities. *Pedagogy, Culture & Society, 15*(1), 55-69. https://doi.org/10.1080/14681360601162212

Taras, M. (2008). Issues of power and equity in two models of Self-assessment. *Teaching in Higher Education, 13*(1), 81-92. http://dx.doi.org/10.1080/13562510701794076

Tashakkori, A., & Creswell, J. W. (2007). Exploring the nature of research questions in mixed methods research. *Journal of Mixed Methods Research, 1*(3), 207-211.

Tashakkori, A., & Teddlie, C. (2003). *Handbook of mixed methods in social and behavioral research.* Thousand Oaks, CA.

Thorndike, E. L. (1913). *The psychology of learning* (Vol. 2). Teachers College, Columbia University.

Thorndike, E. L. (1927). The law of effect. *American Journal of Psychology, 39*(1), 212–222.

Tillema, H. H. (1998). Design and Validity of a Portfolio Instrument for Professional Training. *Studies in Educational Evaluation, 24*(3), 263-78.

Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research, 68*(3), 249-276. http://dx.doi.org/10.3102/00346543068003249

Topping, K. (2003). Self and peer assessment in school and university: Reliability, validity and utility. In Segers, M., Dochy, F., & Cascallar, E. (Eds.). *Optimising new modes of assessment: In search of qualities and standards* (Vol. 1), (pp. 55-

87). Springer Science & Business Media.

Topping, K. (2019). *Using peer assessment to inspire reflection and learning*. Routledge.

Topping, K. (2021). Peer assessment: Channels of operation. *Education Sciences, 11*(3), 91. http://dx.doi.org/10.3390/educsci11030091

Topping, K. J., & Ehly, S. W. (2001). Peer assisted learning: A framework for consultation. *Journal of Educational and Psychological Consultation, 12*(2), 113-132. http://dx.doi.org/10.1207/S1532768XJEPC1202_03

Torrance, H. (1993). Formative Assessment: some theoretical problems and empirical questions. *Cambridge Journal of Education, 23*(3), 333-343. https://doi.org/10.1080/0305764930230310

Torrance, H., & Pryor, J. (1998). Investigating formative assessment: Teaching, learning and assessment in the classroom. McGraw-Hill Education (UK).

Tribble. C. (1996*). Writing*. Oxford University Press.

Truemper, C.M., (2004). Using scoring rubrics to facilitate assessment and evaluation of graduate level nursing students. *Journal of Nursing Education, 43,* (12) 562-4.

Truscott, J., & Hsu, A. Y.-p. (2008). Error correction, revision, and learning. *Journal of Second Language Writing, 17*(4), 292-305. http://dx.doi.org/10.1016/j.jslw.2008.05.003

Truscott, J., (1996). The case against grammar correction in L2 writing classes. *Language Learning 46*(2), 327–369. http://dx.doi.org/10.1111/j.1467-1770.1996.tb01238.x

Turley, E. D., & Gallagher, C. W. (2008). On the" uses" of rubrics: reframing the great rubric debate. *English Journal, 97*(4), 87-92. http://dx.doi.org/10.2307/30047253

Vaismoradi, M., Jones, J., Turunen, H., & Snelgrove, S. (2016). *Theme development in qualitative content analysis and thematic analysis. Journal of Nursing Education and Practice 6*(5).

van Gennip, N., Segers, M., & Tillema, H. H. (2009). Peer assessment for learning from a social perspective: The influence of interpersonal variables and structural features. *Educational Research Review, 4*(1), 41–54. http://dx.doi.org/10.1016/j.edurev.2008.11.002

Värlander, S. (2008). The role of students' emotions in formal feedback situations. *Teaching in Higher Education, 13*(2), 145-156. http://dx.doi.org/10.1080/13562510801923195

Vasileiadou, D., & Karadimitriou, K. (2021). Examining the impact of self-assessment with the use of rubrics on primary school students' performance. *International Journal of Educational Research Open, 2.* http://dx.doi.org/10.1016/j.ijedro.2021.100031

Vygotsky, L. (1978). In M. Cole, V. John-Steiner, S. Scribner, & E. Souberman (Eds.), *Mind in society: The development of higher psychological processes.* Harvard University Press

Vygotsky, L. S. (1987). *The collected works of LS Vygotsky: Problems of the theory and history of psychology* (Vol. 3). Springer Science & Business Media.

Wachwa, G., Schulz, H., & Mann, B. L. (2006). Effects of anonymity and accountability during online peer assessment. In B. L. Mann (Ed.), *Selected styles of web-based educational research* (pp. 300-331). Hershey: Information Science Publishing.

Wang, W. (2014). Students' perceptions of rubric-referenced peer feedback on EFL writing: A longitudinal inquiry. *Assessing Writing, 19,* 80-96. http://dx.doi.org/10.1016/j.asw.2013.11.008

Wang, W. (2016). Using rubrics in student self-assessment: student perceptions in the English as a foreign language writing context. *Assessment & Evaluation in Higher Education, 42*(8), 1280-1292. http://dx.doi.org/10.1080/02602938.2016.1261993

Watson, D. T. (2004). The role of English in the provision of high quality education in the United Arab Emirates [Doctoral dissertation]. University of United Arab Emirates.

Webb, E.J., Campbell, D.T ., Schwartz, R.D. and Sechrest, L. (1966). *Unobtrusive Measures: Nonreactive Research in the Social Sciences*. Rand McNally.

Weigle, S. C. (2002). *Assessing writing*. Cambridge University Press.

Weldon, Fay.2009. On Assessing Creative Writing. *New Writing 6* (3), 168–174. http://dx.doi.org/10.1080/14790720903556734

Wertsch, J. V. (1979). From social interaction to higher psychological processes: A clarification and application of Vygotsky's theory. *Human Development, 22*(1), 1–22. http://dx.doi.org/10.1159/000272425

White, R. (1987). *Writing advanced.* Oxford University Press.

White, R. T. (1994). Conceptual and conceptional change. *Learning and Instruction*, *4*(1), 117-121. https://www.sciencedirect.com/science/article/abs/pii/0959475294900221

Widdowson, H. G. (1995). Discourse analysis: a critical view. Language *and literature*, *4*(3), 157-172. http://dx.doi.org/10.1177/096394709500400301

Wiggins, G. (1989). A true test: Toward more authentic and equitable assessment. *Phi Delta Kappan, 92*(7), 81-93. http://dx.doi.org/10.1177/003172171109200721

Wiggins, G. (1993). *Assessing Student Performance: Exploring the purpose and limits of testing.* Jossey-Bass Publishers.

Wiliam, D. (2007). Keeping learning on track: Formative assessment and the regulation of learning. *Nature, 1*, 20.

Wiliam, D. (2010). An integrative summary of the research literature and implications for a new theory of formative assessment. In Andrade, H. L., & Cizek, G. J. (eds.), *Handbook of formative assessment* (pp. 18-40). Routledge.

Wiliam, D. (2011). *Embedded formative assessment*. Solution Tree Press.

Wiliam, D. (2018). Feedback: At the heart of-but definitely not all of-formative assessment. In Lipnevich, A. A., & Smith, J. K. (Eds.). *The Cambridge handbook of instructional feedback* (pp. 3-28). Cambridge University Press.

Wiliam, D. (2020). Formative assessment, assessment for learning, and all that jazz. https://www.schooleducationgateway.eu/en/pub/viewpoints/experts/formative-assessment-learning.htm

Wiliam, D., & Thompson, M. (2008). Integrating assessment with learning: What will it take to make it work? In C. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning* (pp. 53–84). Lawrence Erlbaum Associates.

Wiliam, D., Lee, C., Harrison, C., & Black, P. (2004). Teachers developing assessment for learning: Impact on student achievement. *Assessment in Education: Principles, Policy & Practice, 11*(1), 49-65. http://dx.doi.org/10.1080/0969594042000208994

Williams, R. L., Carroll, E., & Hautau, B. (2005). Individual accountability in cooperative learning groups at the college level: Differential effects on high, average, and low exam performers. *Journal of Behavioral Education, 14*(3), 167-188. http://dx.doi.org/10.1007/s10864-005-6296-3

Wilson, M. (2007). Why I won't be using rubrics to respond to students' writing. *English Journal, 96*(4), 62-66. http://dx.doi.org/10.2307/30047167

Wolf, D.P. (1989). Portfolio assessment; sampling students' work. *Educational Leadership, 46*(7), 4–10.

Wyatt-Smith, C., & Klenowski, V. (2014). Elements of better assessment for the improvement of learning. In Wyatt-Smith, C., Klenowski, V., & Colbert, P. (Eds.), *Designing assessment for quality learning* (pp. 195-210). Springer.

Wyrick, J. (1999). *Steps to writing well*. Harcout Brace & Company.

Yan, Z., & Brown, G. T. L. (2017). A cyclical self-assessment process: Towards a model of how students engage in self-assessment. *Assessment & Evaluation in Higher Education, 42*(8), 1247-1262. http://dx.doi.org/10.1080/02602938.2016.1260091

Yanqun, H. (2009). L1 influence on L2 writing: Process and product. *CELEA Journal, 32*(2), 3-15.

Yorke, M. (2003). Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher Education, 45*(4), 477-501. https://doi.org/10.1023/a:1023967026413

Yorke, M., & Longden, B. (2004). *Retention and student success in higher education*. McGraw-Hill Education.

Yu, A. (2012). Analysis of the Problems of the Chinese College Students' EFL Classroom Writings. *International Education Studies*, *5*(5), 199-203.

Yu, F. Y. (2012). Any effects of different levels of online user identity revelation? *Journal of Educational Technology and Society,* *15*(1), 64–77. https://www.jstor.org/stable/10.2307/jeductechsoci.15.1.64

Yu, F. Y., & Wu, C. P. (2011). Different identity revelation modes in an online peer-assessment learning environment: Effects on perceptions toward assessors, classroom climate and learning activities. *Computers & Education, 57*(3), 2167–2177. http://dx.doi.org/10.1016/j.compedu.2011.05.012

Zamel, V. (1982). Writing: The process of discovering meaning. *TESOL Quarterly, 16* (2), 195-209). http://dx.doi.org/10.2307/3586792

Zamel, V. (1983). The composing processes of advanced ESL students: Six case studies. *TESOL quarterly, 17*(2), 165-188. http://dx.doi.org/10.2307/3586647

Zeichner, K., & Wray, S. (2001). The teaching portfolio in US teacher education programs: What we know and what we need to know. *Teaching and Teacher Education, 17*(5), 613-621. http://dx.doi.org/10.1016/s0742-051x(01)00017-8

Zemach, D. E., & Islam, C. (2005a). *Paragraph writing: From sentence to paragraph.* Macmillan.

Zemach, D. E., & Rumisek, L. A. (2003). *College writing: From paragraph to essay.* McMillan.

Zhao, Y. (1998). The effects of anonymity on computer-mediated peer review. *International Journal of Educational Telecommunication, 4*(4), 311-345.

Zimmaro, D. M. (2004). Developing grading rubrics. Measurement and Evaluation Center. *Ölçme ve Değerlendirme,* 1-7.

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In Zimmerman, B. J., & Schunk, D. H (Eds.) *Handbook of self-regulation* (pp. 13-39). Academic Press.

Zimmerman, B. J., & Schunk, D. (2011). *Handbook of self-regulation of learning and performance.* Routledge.

Zimmerman,B.J. & Schunk,D.H. (2001). *Self-regulated learning and academic achievement*. Mahwah, NJ, Lawrence Erlbaum Associates.

Zou, M., Kong, D., & Lee, I. (2021). Teacher Engagement with Online Formative Assessment in EFL Writing During COVID-19 Pandemic: The Case of China. *The Asia-Pacific Education Researcher,* 1-12.

APPENDICES

Appendix 1: Pre-study Questionnaire

Appendix 2: Self- and APA of the Descriptive Genre

Appendix 3: Self- and APA of the Narrative Genre

Appendix 4: Rubrics' Description Levels

Appendix 5: Rubrics' Follow-up Sections

Appendix 6: Participants' Sheet to Memorize the Descriptors

Appendix 7: Involving Participants in Managing Classroom Writing-basedAssessment

Appendix 8: Researcher Journal Report

Appendix 9: Post-study Questionnaire for the Experiment Group (Part I) and Post-study Questionnaire for the Experiment Group (Part II)

Appendix 10: Post-study Questionnaire for the Control Group

Appendix 11: Scoring Rubric

Appendix 12: WorkingPortfolio Project

Appendix 13: Participants' Writing Progress in the Intervention Group

Appendix 14: Analysis Methods Used

Appendix 1: Pre-study Questionnaire

Pre-study Questionnaire of Writing Classroom-based Assessment

Dear student,

You are kindly invited to answer the following questions honestly and carefully. All your answers will be considered correct and will be kept strictly confidential.

Language of answering is English. If you find the difficulty to express your ideas in English, you can use French or Arabic language.

Section 01:Background Information

1. Age: ....years old

2. Gender:

☐ Male
☐ Female

3. What was your stream in 3rd year secondary school?

☐ Scientific stream
☐ Literature and philosophy
☐ Literature and languages
☐ Other (Please, precise)

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

4. Is it your choice to study English language at university?

☐ Yes
☐ No

5. What was your mark in English Language in Baccalaureate exam: ..../20

Section 02: Experience with English Writing and Writing Assessment at Secondary Schools

1. Did your lycée teachers of English language ask you to write during your lessons?

☐ Always (Yes)
☐ Sometimes (Yes)
☐ Never (No)

1.1. If **yes**, did your lycée teacher evaluate what you have written?

☐ Yes
☐ No

If **yes**, what did s/he give you as remarks?

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

1.2. If **yes**, did you like the method s/he was using in evaluating your writing?

☐ Yes
☐ No
☐ A little

Justify your answer

....
....

....

....

....

....

....

2. Were you asked to evaluate your **own writing** in secondary school?

☐ Yes
☐ No

If **yes**, how did you self-assess your writing?

....

....

....

....

....

....

....

3. Were you asked to evaluate your **classmates' writing** in secondary school?

☐ Yes
☐ No

If **yes**, how did you peer assess your writing?

....

....

....

....

....

....

Section 03: Attitudes towards English Writing

1. Do you like writing in English?

☐ Yes
☐ No
☐ A little

2. Do you write in your leisure (free time)?

☐ Yes
☐ No
☐ A little

3. In parallel with your university classes, do you take extra courses in English writing in private schools or academic centers/institutions?

☐ Yes
☐ No

4. If you are asked to write in English, do you prefer to do that

☐ Individually
☐ Collaboratively in pairs
☐ Collaboratively in group

5. Do you feel anxious or stressed when you write in English:

☐ Yes
☐ No
☐ A little

6. Do you revise your writing after you finish?

☐ Yes

☐ No

Section 04: Beliefs about English Writing and Writing Assessment

1.What do you think about your level in writing?

☐ Excellent
☐ Good
☐ Average
☐ Poor

2. Do you think you have self-confidence when you write in English?

☐ Yes
☐ No
☐ A little

3. If you don't write well, is that because you think you

☐ cannot
☐ do not want to

Justify your answer

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

4. Do you think you are a good user of grammar?

☐ Yes
☐ No
☐ I don't know

If **not**, what are the difficulties you face in English grammar?

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

5. Do you think you have a rich vocabulary in English?

☐ Yes
☐ No
☐ I don't know

5.1. If **yes**, did you acquire it from (you can tick more than one answer):

☐ Different Readings
☐ Watching English TV programs like BBCW, movies, documentaries, etc.
☐ Watching videos
☐ Different Listening
☐ Others (Please specify)

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

5.2. If **not**, why?

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

6. If I ask you to evaluate your **own writing**, will you be able to do it?

6.1. If **yes**, how?

....

....

....

....

....

....

6.2. If **not**, why?

....

....

....

....

....

....

7. If I ask you to evaluate your **classmate's writing,** will you be able to do it?

☐ Yes
☐ No

7.1. If **yes**, how?

....

....

....

....

....

....

7.2. If **not**, why?

....

....

....

....

....

....

Section 05: Needs and Awareness towards English Writing

1. Do you know how to improve your English writing?

☐ Yes

☐ No

If **yes**, how?

....

....

....

....

....

....

2. Do you need an individual talk with your teacher of writing to expose your problems in writing?

☐ Yes
☐ No

Justify your answer.

....

....

....

....

....

....

3. Marks are better than feedback

☐ Yes

- ☐ No

Justify your answer

....

....

....

....

....

....

....

4. To improve your writing, which you prefer

- ☐ giving marks
- ☐ giving feedback
- ☐ Both

5. Imagine you are the teacher of writing, what do you do to help your students become good writer in English?

....

....

....

....

....

....

....

....

....

6. If you want to say something else in relation to this topic, you are welcome.

....

....

....

....

....

....

....

Thank You for Your Collaboration

Appendix 2: Self- and APA of the Descriptive Genre

Self-assessment of the Descriptive Paragraph *The assessor's Full Name:…………………….*

Anonymous Peer Assessment of the Descriptive Paragraph *The Code……………..*

Paragraph Form

- ☐ Yes
- ☐ No

Handwriting

- ☐ Ex.
- ☐ Gd.
- ☐ Av
- ☐ Pr.

Paragraph Title

- ☐ Ex.
- ☐ Gd.
- ☐ Av
- ☐ Pr.

Mechanics		
Sp.	**Cp.**	**Pn.**
No.	Ex.	Ex.
Rr.	Gd.	Gd.
Fr.	Av.	Av.
Hv.	Pr.	Pr.
Comment……………………….		

Vocabulary		
UV	**KoV**	**FV**
Ex.	Ex.	No.
Gd.	Gd.	Fr.
Av.	Av.	Av.
Pr.	Pr.	Hv.
Comment…………………………		

Language use									
Const.	**S-V agr**	**Tenses**	**UoAr.**	**UoPr.**	**UoPp.**	**Neg.**	**GFW**	**MaS**	**Sce. Prb.**
Ex.	Ex.	Ex.	Ex.	Ex.	Ex.	Ex.	Ex.	Ex.	No
Gd.	Gd.	Gd.	Gd.	Gd.	Gd.	Gd.	Gd.	Gd.	Fr.
Av.	Av.	Av.	Av.	Av.	Av.	Av.	Av.	Av.	Av.
Pr.	Pr.	Pr.	Pr.	Pr.	Pr.	Pr.	Pr.	Pr.	Hv.
Comment..									

Organisation			
SI	**EI**	**CbI**	**SO**
Ex.	Ex.	Ex.	Ex.
Gd.	Gd.	Gd.	Gd.
Av.	Av.	Av.	Av.
Pr.	Pr.	Pr.	Pr.
Comment…………………………..........			

Content			
TS	**SS**	**CS**	**RtT**
Ex.	Ex.	Ex.	Ex.
Gd.	Gd.	Gd.	Gd.
Av.	Av.	Av.	Av.
Pr.	Pr.	Pr.	Pr.
Comment………………….................			

Final Assessment ....

Appendix 3: Self- and APA of the Narrative Genre

Self-assessment of the Narrative Paragraph *The assessor's Full Name:*………………………………….

Anonymous Peer Assessment of the Narrative Paragraph *The Code*…………………………

Paragraph Form

- ☐ Yes
- ☐ No

Handwriting

- ☐ Ex.
- ☐ Gd.
- ☐ Av
- ☐ Pr.

Paragraph Title

- ☐ Ex.
- ☐ Gd.
- ☐ Av
- ☐ Pr.

Mechanics		
Sp.	**Cp.**	**Pn.**
No	Ex.	Ex.
Rr.	Gd.	Gd.
Fr.	Av.	Av.
Hv.	Pr.	Pr.
Comment…………………………………........		

Vocabulary		
UV	**KoV**	**FV**
Ex.	Ex.	No.
Gd.	Gd.	Fr.
Av.	Av.	Av.
Pr.	Pr.	Hv.
Comment……………………………….......		

Language use									
Const.	**S-V agr**	**Tenses**	**UoAr.**	**UoPr.**	**UoPp.**	**Neg.**	**GFW**	**MaS**	**Sce Prb**
Ex.	Ex.	Ex.	Ex.	Ex.	Ex.	Ex.	Ex.	Ex.	No
Gd.	Gd.	Gd.	Gd.	Gd.	Gd.	Gd.	Gd.	Gd.	Fr.
Av.	Av.	Av.	Av.	Av.	Av.	Av.	Av.	Av.	Av.
Pr.	Pr.	Pr.	Pr.	Pr.	Pr.	Pr.	Pr.	Pr.	Hv.
Comment..									

Organisation			
SI	**EI**	**CbI**	**TO**
Ex.	Ex.	Ex.	Ex.
Gd.	Gd.	Gd.	Gd.
Av.	Av.	Av.	Av.
Pr.	Pr.	Pr.	Pr.
Comment………………………………...........			

Content			
TS	**SS**	**CS**	**RtT**
Ex.	Ex.	Ex.	Ex.
Gd.	Gd.	Gd.	Gd.
Av.	Av.	Av.	Av.
Pr.	Pr.	Pr.	Pr.
Comment…………………….................…			

Final Assessment ....

Appendix 4: Rubrics' Description Levels

Rubrics' Description Levels

Dear student,

Presented below are the description levels of the criteria of the rubrics you are using to assess your own and your classmate's writings. Try to read and follow carefully these descriptors while using the rubrics.

Never hesitate to ask questions if you have any.

For the counting technique-0123:

a) 0 means there is **no** mistake for excellent.
b) 1 means there is **one** mistake for good.
c) 2 means there are **two** mistakes for average.
d) 3 means there are **three** or more mistakes for poor.

1-Paragraph Form

1. **Yes:** the paragraph must be both indented and written as a block.
2. **No:** the paragraph is not indented, not written as a block, or both are missed i.e. at least condition one is missed or both.

2-Handwriting

1. **Excellent:** nice and readable.
2. **Good:** simple and readable.
3. **Average:** readable, but not nice.
4. **Poor:** hardly readable and messy.

3-Paragraph Title

1. **Excellent:** it has a metaphoric meaning and topic-relevant.
2. **Good:** target the main idea and linguistically well structured.
3. **Average:** direct and but not that expressive.
4. **Poor:** very general and may include linguistic inaccuracies, irrelevant, or no title at all.

4-Mechanics

✓ **Sp: Spelling mistakes**Counting technique (0123-technique)

1. **Excellent:** no spelling mistakes
2. **Good:** one spelling mistake
3. **Average:** two spelling mistakes
4. **Poor:** three or more spelling mistakes

✓ **Cp: Capitalisation** Counting technique (0123-technique)

1. **Excellent:**no problem with capitalization
2. **Good:** one word is not capitalized
3. **Average:** two words are not capitalized
4. **Poor:** three or more words are not capitalized

✓**Pn: Punctuation**Counting technique (0123-technique)

1. **Excellent:**no problem with punctuation marks.

2. **Good:** one mistake in punctuation marks.
3. **Average:** two mistakes in punctuation marks.
4. **Poor:** three or more mistakes in punctuation marks.

5-Vocabulary

- ✓ **UV: Use of vocabulary** Counting technique (0123-technique)
  1. **Excellent:** appropriate contextualized meaning of all words
  2. **Good:** one word has not an appropriate contextualized meaning.
  3. **Average:** two words have not an appropriate contextualized meaning.
  4. **Poor:** three words have not an appropriate contextualized meaning.
- ✓ **KoV: Knowledge of vocabulary** (In self-assessment, circle the new words you have recently acquired and used in your written piece. In peer assessment, circle the new words you didn't know their meaning using your pencil).
  1. **Excellent:**if you find 5 words in your peer writing (in peer assessment) that you don't know their meanings oryou use 5 new words in your own writing (in self-assessment)
  2. **Good:** if you find/ use 4 new words you don't know their meaning
  3. **Average:** if you find/ use 3 new words you don't know their meaning
  4. **Poor:**if you find/ acquire just 2 or less new words you don't know their meaning
- ✓ **FV: Foreign Vocabulary**Counting technique (0123-technique)
  1. **Excellent:** no use of foreign vocabulary
  2. **Good:** one foreign word is found.
  3. **Average:** two foreign words are found.
  4. **Poor:** three or more foreign words are found.

6-Language Use

- ✓ **Const.: Constructions**
  1. **Excellent**: if you use correctly all different types of sentences: compound-complex, complex, compound, and simple sentences.
  2. **Good:** if you use correctly just complex, compound, and simple sentences.
  3. **Average**: if you use correctly just compound and simple sentences.
  4. **Poor:** if you use correctly only simple sentences.
- ✓ **S-V agr : subject-verb agreement**Counting technique (0123-technique)
  1. **Excellent:** all sub-verb agreement are correct
  2. **Good:** one sub-verb agreement is not correct

Appendix 4: Rubrics' Description Levels

3. **Average:** two sub-verb agreements are not correct.
4. **Poor:** three or more sub-verb agreements are not correct.

- ✓ **UoAr: Use of articles**Counting technique(0123-technique)
  1. **Excellent:** all articles are used correctly.
  2. **Good:** one article is not used correctly
  3. **Average:** two articles are not used correctly.
  4. **Poor:** three or more articles are not used correctly.
- ✓ **UoPr: Use of pronouns**Counting technique(0123-technique)
  1. **Excellent:** all pronouns are used correctly
  2. **Good:** one pronoun is not used correctly
  3. **Average:** two pronouns are not used correctly.
  4. **Poor:** three or more pronouns are not used correctly
- ✓ **UoPp: Use of prepositions**Counting technique (0123-technique)
  1. **Excellent:** all prepositions are used correctly
  2. **Good:** one prepositions is not used correctly
  3. **Average:** two prepositions are not used correctly.
  4. **Poor:** three or more prepositions are not used correctly
- ✓ **Neg: Negations (Appropriate use, if any, of negative forms)**Counting technique(0123-technique)
  1. **Excellent:** all negative forms are correct
  2. **Good:** one negative form is not correct
  3. **Average:** two negative forms are not correct.
  4. **Poor:** three or more negative forms are not correct.
- ✓ **GFW: grammatical functions of words**Counting technique(0123-technique)
  1. **Excellent:** all words are grammatically correct.
  2. **Good:** one word is not grammatically correct.
  3. **Average:** two pronouns are not used correctly.
  4. **Poor:** three or pronouns are not used correctly.
- ✓ **MaS: meaning at sentence level** (meaningful and context-related sentences) . Counting technique (0123-technique)
  - ✓ **Excellent:** all sentences are meaningful.
  - ✓ **Good:** one sentence is not meaningful
  - ✓ **Average:** two sentences are not meaningful.

Appendix 4: Rubrics' Description Levels

- ✓ **Poor:** three or more sentences are not meaningful.

✓ **Sce Prb: Sentence problems (All problems were included: fragment, run-on, comma splice, fused, choppy, stringy, and non-parallel structure)** Counting technique (Whatever the problem is. For example, if you have one choppy sentence and one stringy, that means there are 2 problems and this is average. If you have one run-on sentence, one fragment, one non-parallel means that the mistakes are three and this means poor.)

1. **Excellent:** no sentence problems are found.
2. **Good:** one sentence problem is found.
3. **Average:** two sentence problems are found.
4. **Poor:** three or more sentence problems are found.

7-Organization

✓ **SI: Stated Ideas**

1. **Excellent**: all main and expressive ideas are listed.
2. **Good**: listing relevant ideas to the context.
3. **Average**: listing just secondary ideas which are not strong enough.
4. **Poor** : just weak ideas poorly developed

✓ **EI: Expressive Ideas**

1. **Excellent**: strong ideas explained fully.
2. **Good**: more or less extensive explanation.
3. **Average**: lacking explanations.
4. **Poor**: practically there are no explanations.

✓ **CbI: connection between ideas**

1. **Excellent:** perfect use of appropriate connectors respecting the genre and smooth transition playing on linguistic constructions.
   - Descriptive genre: using connectors specific to the genre to make the image clear.
   - Narrative genre: using appropriate connectors of times
2. **Good:** rarely missed connectors when it is needed, may be one connector, but cohesion and coherence is still maintained.
3. **Average:** lacking some connectors and weaknesses in manipulating the language, may be two connectors which shake the cohesion and the coherence.
4. **Poor:** practically no connectors and poor linguistic constructions, three or more, which lead to very weak cohesion and coherence.

✓ **SO: Order of importance/ Space order**

Appendix 4: Rubrics' Description Levels

1. **Excellent:** Respecting the space order creating an image of the thing/person described. Clear focus and harmony in description
2. **Good:** the image is still clear, but it lacks some minor details.
3. **Average:** the image is not clear.
4. **Poor:** no clear description, just random ideas.

- ✓ **TO: Time order**
  1. **Excellent:** respecting the structure of telling a story: beginning climax and the end
  2. **Good:** respecting the structure of the story with little missing details.
  3. **Average:** the story is not that clear. Telling the story without respecting the structure. It is messy
  4. **Poor:** there is no structure of a story.

8-Content

- ✓ **TS: Topic sentence**
  1. **Excellent:** very clear topic with very clear controlling idea
     - Descriptive genre, the dominant impression(controlling idea) is apparent and well expressed
     - Narrative genre, dominant feeling (controlling idea)is clearly expressed
  2. **Good:** the topic is more or less clearly expressed but clear expressive controlling idea.
  3. **Average:** the controlling idea is not expressed fully.
  4. **Poor:** no structure of a topic sentence or does not exist at all. Just an empty introductory sentence.
- ✓ **SS: Supporting sentences**
  1. **Excellent**: topic is fully developed with enough quality ideas respecting the features of the genre.
  2. **Good:** topic is developed enough with slight missing details.
  3. **Average:** lacks development that affect the overall meaning
  4. **Poor:** practically there is no development of the topic.
- ✓ **CS: Concluding sentence**
  1. **Excellent:**
     - Descriptive genre: It restates the description in brief with a final thought.
     - Narrative genre: It gives an impression or a lesson through telling that story
  2. **Good:**clear and understood concluding sentence.
  3. **Average:**Unclear concluding sentence with ambiguous meaning.
  4. **Poor:** Poor concluding sentence, irrelevant, or missed at all.
- ✓ **RtT: Relevance to the topic.**
  1. **Excellent:** all ideas are expressed perfectly and related to the topic
  2. **Good:** one idea is not related
  3. **Average:** two ideas are not relevant
  4. **Poor:** three or more are not relevant

Appendix 5: Rubrics’ Follow-up Sections

Self-assessment ***The Assessors’ Full Name....................................***

Message from the **student writer** and **user of the rubric** to the teacher

1. My writing

1.1.I am satisfied (what and why)

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

1.2. I am not satisfied (what and why)

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

2.My use of the rubric

2.1. I am satisfied (what and why)

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

2.2. I am not satisfied (what and why)

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

Appendix 5: Rubrics' Follow-up Sections

Anonymous Peer Assessment ***The Code…………………***

1. Message from the **student user of the rubric**

1.1. What do you think about your classmate's writing?

……
……
……
……
……
……

1.2. Did the rubric help you to assess your classmate's writing? If **yes**, how? If **no,** why?

……
……
……
……
……
……

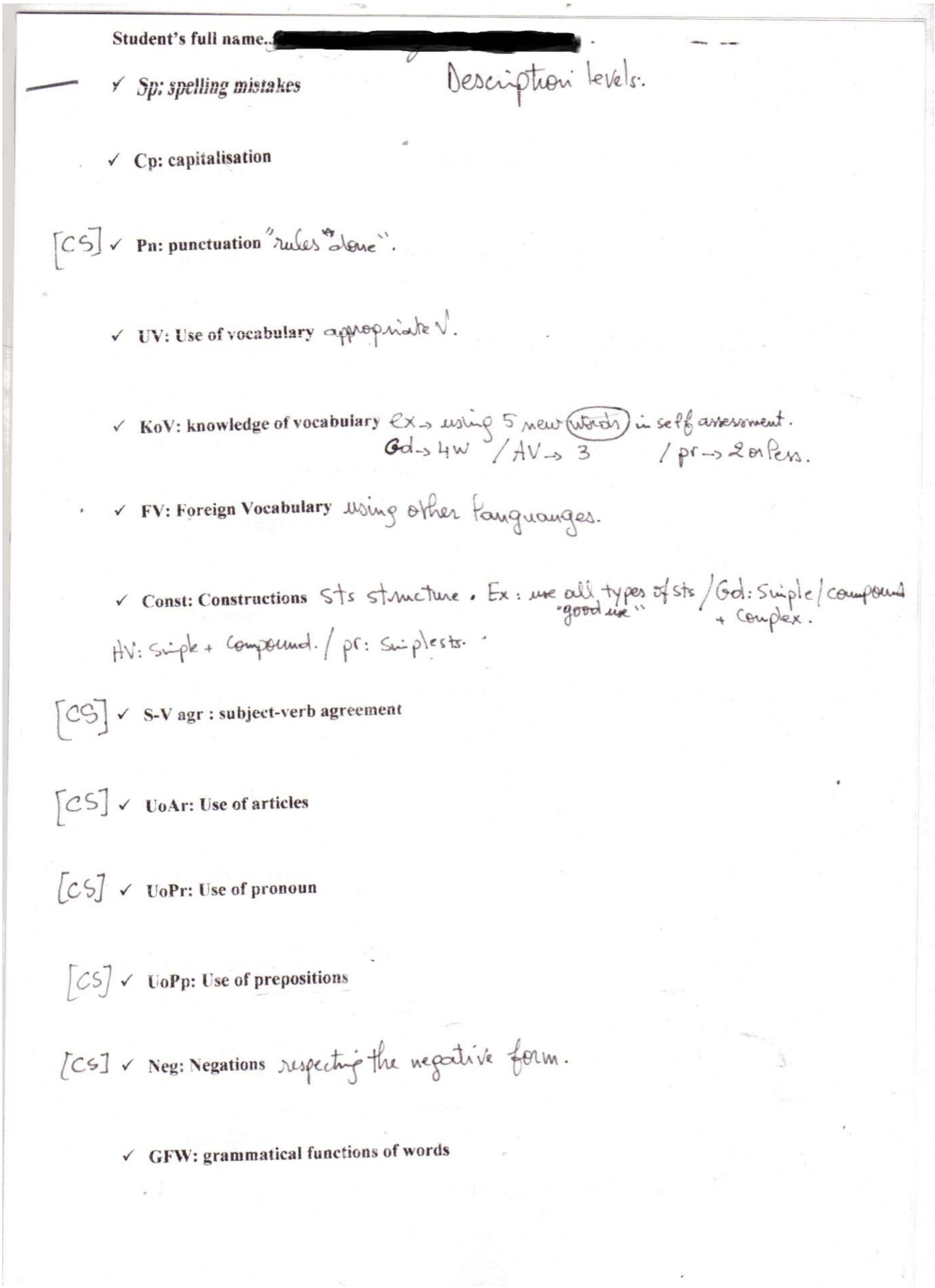
2. Message from the **student writer**

What do you think about the **assessment of your classmate of your writing using the rubric ?**

……
……
……
……
……
……
……
……
……
……
……

Appendix 6: Participants’ Memorization Sheet

Participants’ Sheet to Memorize the Descriptors



Student’s full name..[redacted]

✓ *Sp: spelling mistakes* Description levels.

✓ **Cp: capitalisation**

[CS] ✓ **Pn: punctuation** "rules alone".

✓ **UV: Use of vocabulary** appropriate V.

✓ **KoV: knowledge of vocabulary** Ex -> using 5 new (words) in self assessment.
Gd -> 4W / AV -> 3 / pr -> 2 or less.

✓ **FV: Foreign Vocabulary** using other languanges.

✓ **Const: Constructions** Sts structure. Ex: use all types of sts "good use" / Gd: simple/compound + complex.
AV: simple + compound. / pr: simple sts.

[CS] ✓ **S-V agr : subject-verb agreement**

[CS] ✓ **UoAr: Use of articles**

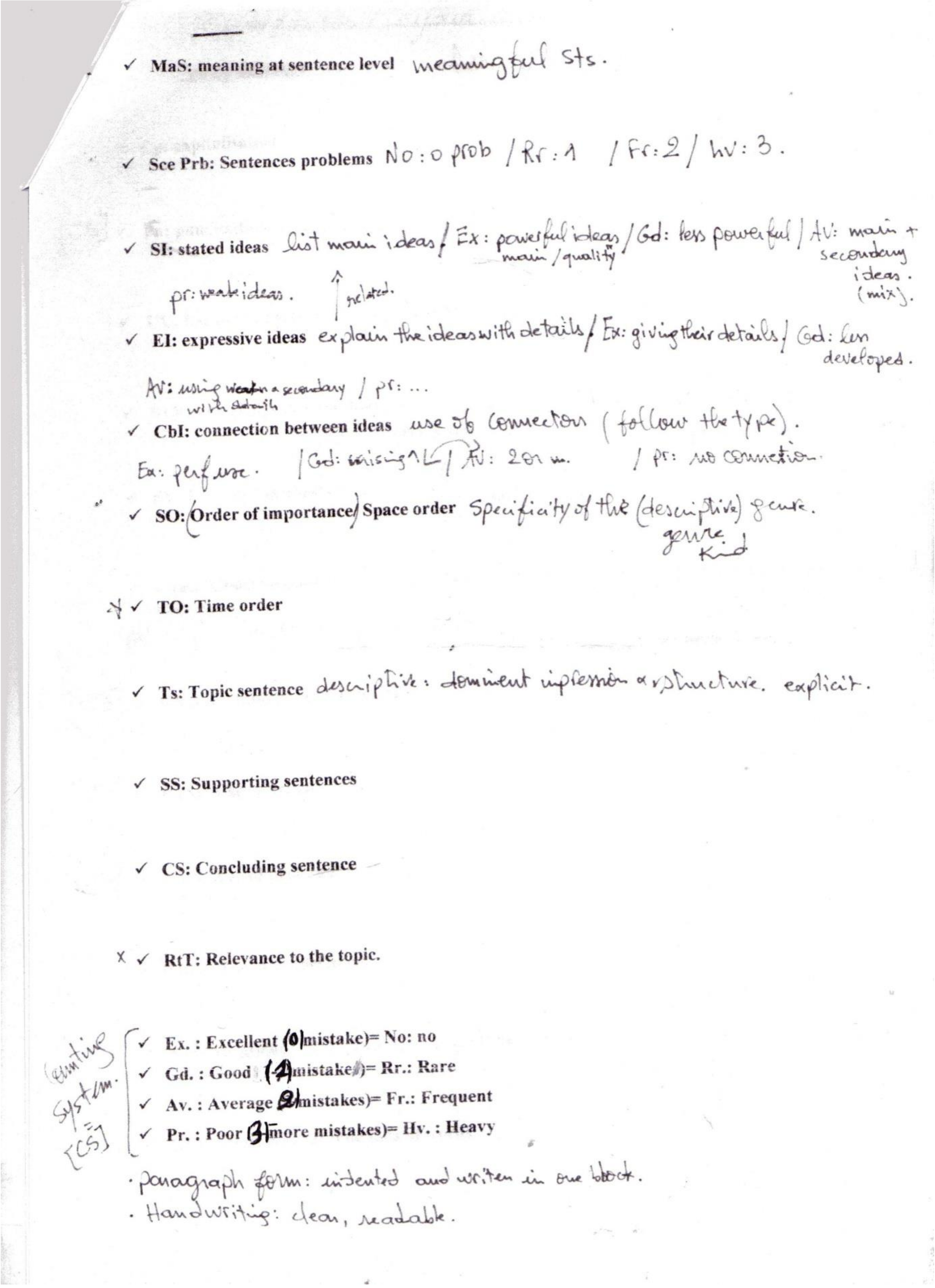
[CS] ✓ **UoPr: Use of pronoun**

[CS] ✓ **UoPp: Use of prepositions**

[CS] ✓ **Neg: Negations** respecting the negative form.

✓ **GFW: grammatical functions of words**

Appendix 6: Participants' Memorization Sheet



✓ **MaS: meaning at sentence level** meaningful Sts.

✓ **Sce Prb: Sentences problems** No : 0 prob / Rr : 1 / Fr : 2 / hv : 3.

✓ **SI: stated ideas** list main ideas / Ex : powerful ideas main / quality / Gd: less powerful / AV: main + secondary ideas. (mix).
pr: weak ideas. ↑ related.

✓ **EI: expressive ideas** explain the ideas with details / Ex: giving their details / Gd: less developed.
AV: using ~~weak~~ a secondary with details / pr: ...

✓ **CbI: connection between ideas** use of connectors (follow the type).
Ex: perf use. / Gd: missing ^L / AV: 2 or m. / pr: no connection.

✓ **SO: Order of importance/ Space order** Specificity of the (descriptive) genre. genre Kind

✗ ✓ **TO: Time order**

✓ **Ts: Topic sentence** descriptive: dominent impression or structure. explicit.

✓ **SS: Supporting sentences**

✓ **CS: Concluding sentence**

x ✓ **RtT: Relevance to the topic.**

Counting system. = [CS]
- ✓ **Ex. : Excellent (0 mistake)= No: no**
- ✓ **Gd. : Good (1 mistake)= Rr.: Rare**
- ✓ **Av. : Average (2 mistakes)= Fr.: Frequent**
- ✓ **Pr. : Poor (3 more mistakes)= Hv. : Heavy**

· paragraph form: intented and writen in one block.
· Handwriting: clean, readable.

Appendix 7: Involving Participants in Managing Classroom Writing-based Assessment

Involving Participants in Managing Writing Classroom-Based Assessment

Dear student,

To participate and contribute to your own progress in writing through being engaged in the writing assessment classroom, please answerhonestly and seriously the questions below.

Your answers will be very helpful for the teacher to plan the activities to assess your and your classmates' writings to improve your writing.

1. How much time do you think you need to write one (1) paragraph?

- ☐ 20 minutes
- ☐ 30 minutes
- ☐ 40 minutes
- ☐ More, please specify...

2. What types of paragraph you prefer to write?

  - ☐ Descriptive and narrative
  - ☐ Descriptive and opinion
  - ☐ Narrative and opinion
  - ☐ Others, please specify (write the types in pairs like above

    ....

3. Do you think that putting marks like 04/20, 12/20, etc on your writing is better than providing written remarks like good paragraph, poor paragraph, etc?

- ☐ Yes
- ☐ No

Justify your answer……………………………………………………………….............……………
…….……………………………………………………………………………..………………
…………………...………………………………………………………………………………
……………………………………………………………………………...............

4. Which do you prefer?

☐ Identified peer assessment

☐ Anonymous peer assessment

Justify your answer…………………………………………………………………………….....

.………………….……………………………………………………………………………

……………………………………...………………………………………………………

…………………………………………………………………………………….............

5. If your classmate reads your piece of writing and provides you with remarks, will you accept them?

☐ Yes, I do like it and I do take into account what s/he said.

☐ Yes, but I don't like the issue and I will never take into account what s/he said.

☐ No, I don't like the issue but after all I will come back revising what I had done.

☐ No, and I will not accept it at all.

6. Which of these two suggestions you think will help you to improve your writing?

☐ Doing self-assessment then peer assessment then self assessment and so on i.e. consecutively

☐ Doing self-assessment for a number of times (for example 3 times) then peer assessment for a period of time i.e. non-consecutively.

Justify your answer ………………………………………………………………………..

7.Which do you prefer?

☐ Doing self- and peer assessment in descriptive genre, then doing the same thing for the narrative genre i.e. no mixing genres

☐ Mixing both genres

8. a. How many times do you think you need to perform **self-assessment** to write effectively the **descriptive** paragraph?……..…………… ………………………………………………

b. How many times do you think you need to perform **self-assessment** to write effectively the **narrative** paragraph?………… ………………… ………………………………………..

9. a. How many times do you think you need to perform **peer assessment** to write effectively the **descriptive** paragraph?………………………………………………………………….

b. How many times do you think you need to perform **peer assessment** to write effectively the **narrative** paragraph…………………………………………………

10. a. How many **uses of the rubric** do you think you need to write effectively the **descriptive paragraph?**………………………… ….……… ……………………………..

b. How many **uses of the rubric** doyou think you need to write effectively the **Narrative paragraph?**..............……………………………………………………………………..

11**.** Can you write on any topic?

☐ Yes
☐ No
☐ Even though I have information about the topic, I cannot write successfully a paragraph.

If **not**, what are the topics which motivate you to write better?...
....

12. Do you want to know who assessed your writings in anonymous peer assessment?

☐ Yes, after each assessment I want to see who assessed my writing.
☐ Yes, but after we finish all the peer assessment sessions.
☐ Even after we finish all the assessment, I don't want to know who assessed my writing.

Justify your answer

………………………………………………………….............……………………………
………….………………………………………………………………………….…….……
………………………...………………………………………………………………………

13. If you have other ideas about the topic, you can share them.

……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………......
……………………………………………………………………………………………

Thank You for Your Collaboration

Appendix 8: Researcher Journal Report

Researcher Journal Report

What is reported below is purely for the intervention phase where the participants are involved in 12 times doing self and anonymous peer assessment for two genres: the descriptive and the narrative. The major interest covered in this observation journal report are time devoted for participant writing, and doing self-assessment and anonymous peer assessment, and reporting the individual inquiries about rubric use and the overall classroom assessment. This is to examine their responsiveness towards their engagement in classroom assessment.

1. Participant Writing Time

Genre	**Descriptive**											
NoT*	**1**				**2**				**3**			
Time Interval	**]60-45]**	**]45-30]**	**]30-15]**	**]15-0]**	**]60-45]**	**]45-30]**	**]30-15]**	**]15-0]**	**]60-45]**	**]45-30]**	**]30-15]**	**]15-0]**
N**												

***NoT:** Number of times

****N:**Total number of the participants

Genre	**Narrative**											
NoT	**1**				**2**				**3**			
Time Interval	**]60-45]**	**]45-30]**	**]30-15]**	**]15-0]**	**]60-45]**	**]45-30]**	**]30-15]**	**]15-0]**	**]60-45]**	**]45-30]**	**]30-15]**	**]15-0]**
N												

2. Time Devoted to Self-assessment

Genre	**Descriptive**											
NoT	**1**				**2**				**3**			
Interval Time	**]60-45]**	**]45-30]**	**]30-15]**	**]15-0]**	**]60-45]**	**]45-30]**	**]30-15]**	**]15-0]**	**]60-45]**	**]45-30]**	**]30-15]**	**]15-0]**
N												

Genre	**Narrative**											
NoT	**1**				**2**				**3**			
Time Interval	**]60-45]**	**]45-30]**	**]30-15]**	**]15-0]**	**]60-45]**	**]45-30]**	**]30-15]**	**]15-0]**	**]60-45]**	**]45-30]**	**]30-15]**	**]15-0]**
N												

3. Time Devoted to Anonymous Peer Assessment

Genre	**Descriptive**											
NoT	**1**				**2**				**3**			
Time Interval	**]60-45]**	**]45-30]**	**]30-15]**	**]15-0]**	**]60-45]**	**]45-30]**	**]30-15]**	**]15-0]**	**]60-45]**	**]45-30]**	**]30-15]**	**]15-0]**
N												

Genre	**Narrative**											
NoT	**1**				**2**				**3**			
Time Interval	**]60-45]**	**]45-30]**	**]30-15]**	**]15-0]**	**]60-45]**	**]45-30]**	**]30-15]**	**]15-0]**	**]60-45]**	**]45-30]**	**]30-15]**	**]15-0]**
N												

4. Individual Inquiries about Writing Rubric Use

4.1 Participants' Individual Inquiries about Descriptors

To help the participants memorize the description levels for running assessment smoothly and easily, the researcher had prepared for them a sheet without any descriptors but only the name of the criteria and sub-criteria. The idea was helping participants to write or summarize the descriptors in the way they help them to fasten their understanding and memorization. This sheet was kept to participant's request while doing the assessment; either to ask for it or not.

Performance	**Assessing descriptive Genre**						**Assessing Narrative Genre**					
NoT	**1(s)***	**2(p)****	**3(s)**	**4(p)**	**5(s)**	**6(p)**	**1(s)**	**2(p)**	**3(s)**	**4(p)**	**5(s)**	**6(p)**
N (Yes)												

***(s):** Self-assessment

****(p):** Peer assessment, which is anonymous

Questions Asked

P1:..

....

....

P2:..

....

....

P3:..

....

....

P4:..

....

....

P5:..

....

....

P6:..

....

....

P7:..

....

....

P8:..

....

....

P9:..

....

....

P10:……………………………………………………………………………………………………

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

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

P11:……………………………………………………………………………………………………

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

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

P12:……………………………………………………………………………………………………

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

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

4.2 Participants' Individual Inquiries about Rubrics' Follow-up Sections

Performance	**Assessing Descriptive Genre**						**Assessing Narrative Genre**					
NoT	**1(s)**	**2(p)**	**3(s)**	**4(p)**	**5(s)**	**6(p)**	**1(s)**	**2(p)**	**3(s)**	**4(p)**	**5(s)**	**6(p)**
N(Yes)												

Questions Asked

P1:……………………………………………………………………………………………………

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

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

P2:……………………………………………………………………………………………………

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

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

P3:……………………………………………………………………………………………………

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

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

P4:……………………………………………………………………………………………………

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

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

P5:……………………………………………………………………………………………………

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

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

P6:...

....

....

P7:...

....

....

P8:...

....

....

P9:...

....

....

P10:...

....

....

P11:...

....

....

P12:...

....

....

4.3 Participants'Individual Inquiries about Delivering Peer feedback

Performance	**Assessing Descriptive Genre**			**Assessing Narrative Genre**		
NoT	**1(p)**	**2(p)**	**3(p)**	**1(p)**	**2(p)**	**3(p)**
N (Yes)						

Questions Asked

P1:...

....

....

P2:...

....

....

P3:...

....

....

P4:...

....

....

P5:...

....

....

P6:...

....

....

5. Participants' Individual Inquiries about Overall Classroom Assessment

Performance	**Assessing descriptive Genre**						**Assessing Narrative Genre**					
NoT	1(s)	2(p)	3(s)	4(p)	5(s)	6(p)	1(s)	2(p)	3(s)	4(p)	5(s)	6(p)
N(Yes)												

Questions Asked

P1:...

....

....

P2:...

....

....

P3:...

....

....

P4:………

……

……

P5:………

……

……

P6:………

……

……

P7:………

……

……

P8:………

……

……

P9:………

……

……

P10:……

……

……

P11:……

……

……

P12:……

……

……

Appendix 9: Post-study Questionnaire (Part I)

Writing Classroom-based Assessment (Part-I)

Dear student,

You are kindly invited to answer the following questions honestly and carefully. All your answers will be considered correct and will be kept strictly confidential.

Section 01: Rubrics' Presentation and Use

1. What do you think about the **form** of the rubrics?

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

2. What do you think about the **organization** of the rubrics?

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

3. What do you think about the **content** of the rubrics?

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

4. What do you think about the **description levels** of the rubrics, were they easy to apply on your writing pieces?

- ☐ Yes
- ☐ No

Justify your choice (Why)

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

5. What do you think about the **function** of the rubrics?

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

6. While using the rubrics, have you used them easily?

- ☐ Yes
- ☐ No

If **yes**, after how many uses did you become familiar with them?

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

7. Did you like the idea of using rubrics?

- ☐ Yes
- ☐ No

Justify your answer (Why)

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

8. If you will be given a second chance to use rubrics, will you like to do that?

- ☐ Yes
- ☐ No

Justify your choice (Why)

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

9. Did you like the scoring rubric?

- ☐ Yes
- ☐ No

Justify your choice (Why)

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

10. If you used the scoring rubric from the beginning, would this impact you to assess objectively your writings in self-assessment?

- ☐ Yes
- ☐ No

Justify your choice (Why)

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

11. If you used the scoring rubric from the beginning, would this impact you to assess objectively your classmate’s writings in anonymous peer assessment?

- ☐ Yes
- ☐ No

Justify your choice (Why)

....

....

....

....

....

....

....

Section 02: Self-assessment Performance

1. Did you like the idea of self-assessment?

- ☐ Yes
- ☐ No
- ☐ A little

2. How did you find self-assessment?

………………………………………………

………………………………………………

………………………………………………

………………………………………………

………………………………………………

…………………………………......................

3. In your opinion, could self-assessment take place without using the rubrics?

- ☐ Yes
- ☐ No

Justify your choice (Why)

....

....

....

....

....

....

....

4. What do you think about the follow-up section associated with the rubrics in self assessment (I am/am not satisfied about my writing & I am/am not satisfied about my use of the rubric)?

....

....

....

....

....

....

....

5. In general, while performing self-assessment, do you think you were

- ☐ very subjective
- ☐ subjective
- ☐ very objective
- ☐ Objective
- ☐ Other

………………………………………………

………………………………………………

………………………………………………

6. If you will be given a second chance to perform self-assessment, will you like to do that?

- ☐ Yes
- ☐ No

Section 03: Anonymous Peer Assessment Performance

Appendix 9: Post-study Questionnaire (Part I)

1. Did you like the idea of anonymous peer assessment?

- ☐ Yes
- ☐ No
- ☐ A little

2. How did you find anonymous peer assessment?

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

3. In your opinion, could anonymous peer assessment take place without using the rubrics?

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

4. What do you think about the follow-up section associated with the rubrics in anonymous peer assessment (message from student user of the rubric and message from student's writer)?

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

5. In general, while doing anonymous peer assessment, do you think you were

- ☐ Very subjective
- ☐ Subjective
- ☐ Very objective
- ☐ Objective
- ☐ Other, please specify

………………...
....
....

6. In general, what can you say about the assessments you received from your classmates?

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

7. In general, what can you say about the remarks or the feedback you received from your classmates?

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

Appendix 9: Post-study Questionnaire (Part I)

....

....

....

8. In general, did you learn from your peers' writings in anonymous peer assessment?

☐ Yes
☐ No

If yes, what did you learn from them?

....

....

....

....

....

....

....

9. If you will be given a second chance to perform anonymous peer assessment, will you like to do that?

☐ Yes
☐ No

Thank you very much for your cooperation

Appendix 9: Post-study Questionnaire (Part II)

Writing Classroom-based Assessment (Part-II)

Dear student,

You are kindly invited to answer the following questions honestly and carefully. All your answers will be considered correct and will be kept strictly confidential.

Section 01: Classroom Assessment

1. Which did you like more?

- ☐ Self-assessment using the rubrics
- ☐ Anonymous peer assessment using the rubrics
- ☐ Both
- ☐ Neither the first, nor the second

Justify your choice (Why)

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

2. Did you like the idea of giving your opinions each time before doing different assessments?

- ☐ Yes
- ☐ No

Justify your choice (Why)

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

3. Was the period of time sufficient to practise your writing using the rubrics?

- ☐ Yes for both, descriptive and narrative genres
- ☐ No for both
- ☐ Yes for the descriptive genre and No for the narrative one
- ☐ No for the descriptive genre and Yes for the narrative one

4. Did you take into account teacher feedback while using rubrics and doing the assessments?

- ☐ Yes
- ☐ No

5. How did you find your experience when you participated and took part in assessing your writing?

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

Appendix 9: Post-study Questionnaire (Part II)

…………………………………………………

…………………………………………………

…………………………………………………

…………………………………………………

…………………………………………………

…………………………………………………

Section 02: Student Writing Progress

1. Do you think that your writing improved?

- ☐ Yes
- ☐ No

1.1. What kind of progress you did

- ☐ Quick progress in both genres
- ☐ Average progress in both
- ☐ Slow progress in both
- ☐ Others, please specify (like slow progress in descriptive and quick in narrative, etc).......................................

....

1.2. How did you achieve this improvement?

....

....

....

....

....

....

....

1.3. This progress was it achieved thanks to

- ☐ Self-assessment
- ☐ Anonymous peer assessment
- ☐ Rubrics
- ☐ All of them
- ☐ None of them

2. Are you aware of the weaknesses, if any, you still have in your writing?

- ☐ Yes
- ☐ No

If **yes**, what are they?

....

....

....

....

....

3. Have you gained any strong points in relation to your writing?

- ☐ Yes
- ☐ No

If **yes**, what are they?

....

....

....

....

....

4. Did you gain self-confidence in your writing in English?

- ☐ Yes
- ☐ No

Appendix 9: Post-study Questionnaire (Part II)

☐ A little

4.1. If you answer by **yes** or**a little**, in which genre you gained this confidence

☐ Narrative
☐ Descriptive
☐ Both

4.2. What did help you to gain this self-confidence?

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

5. Do you think that you still need teacher guidance to write?

☐ Yes
☐ No
☐ A little

6. Are you able to write other types of paragraphs like comparison/contrast, cause/effect, and so on?

☐ Yes
☐ No

Justify your choice (Why)

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

7. Do you revise what you have written?

☐ Yes
☐ No

8.Do you think that you are developing lengthy paragraph compared to first ones?

☐ Yes
☐ No

9. Are you able to compare your writing in the beginning and now?

☐ Yes
☐ No

In the beginning
....
....
....
....
....
....
....

Now...
....
....
....
....
....
....
....

Thank you very much for your cooperation

Appendix 10: Post-study Questionnaire for the Control Group

Writing Classroom

Dear student,

You are kindly invited to answer the following questions honestly and carefully. All your answers will be considered correct and will be kept strictly confidential.

1. Do you think that your writing improved?

☐ Yes
☐ No

1.2. If **yes**, how did you achieve this improvement?

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

1.2. If **no**, why did not you improve your writing?

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

2. Do you feel motivated to improve your writing?

☐ Yes
☐ No
☐ A little

3. Do you still have weaknessesin your writing?

☐ Yes
☐ No

If **yes**, what are they?

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

4. Have you gained any strong points in relation to your writing?

☐ Yes
☐ No

If **yes**, what are they?

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

5. Did you gain self-confidence in your writing in English?

- ☐ Yes
- ☐ No
- ☐ A little

If **yes** or **a little,** what did help you to gain this self-confidence?

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

6. Do you feel stressed or anxious when you write?

- ☐ Yes
- ☐ No
- ☐ A little

7. Do you think that you still need teacher guidance to write?

- ☐ Yes
- ☐ No
- ☐ A little

8. Do you revise what you have written?

- ☐ Yes
- ☐ No

9. Is there a difference between your writings in the beginning and now?

- ☐ Yes
- ☐ No

If **yes**, try to compare between them

In the beginning

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

Now...
....
....
....
....
....
....
....
....
....

Thank you very much for your cooperation

Appendix 11: Scoring Rubric

……/37

Scoring Rubric of the Des./Nar. Paragraph

Student's Full Name…………………………..

Paragraph Form

☐ Yes
☐ No

Handwriting

☐ Ex.
☐ Gd.
☐ Av
☐ Pr.

Paragraph Title

☐ Ex.
☐ Gd.
☐ Av
☐ Pr.

Mechanics/4		
Sp./1	**Cp./1**	**Pn./2**
No.(1.00)	Ex. (1.00)	Ex. (2.00)
Rr.(0.75)	Gd.(0.75)	Gd.(1.50)
Fr.(0.50)	Av.(0.50)	Av.(1.00)
Hv.(0.25)	Pr.(0.25)	Pr.(0.50)
Comment………………………./4		

Vocabulary/6		
UV/2	**KoV/2**	**FV/2**
Ex. (2.00)	Ex. (2.00)	No.(2.00)
Gd.(1.50)	Gd.(1.50)	Fr.(1.50)
Av.(1.00)	Av.(1.00)	Av.(1.00)
Pr.(0.50)	Pr.(0.50)	Hv.(0.50)
Comment…………………………./6		

Language use/ 14									
Const./2	**S-V agr./1**	**Tenses /2**	**UoAr. /1**	**UoPr. /1**	**UoPp. /1**	**Neg. /1**	**GFW /2**	**MaS /1**	**Stce. Prb./2**
Ex. (2.00)	Ex. (1.00)	Ex. (2.00)	Ex. (1.00)	Ex. (1.00)	Ex. (1.00)	Ex. (1.00)	Ex. (2.00)	Ex. (1.00)	No.(2.00)
Gd.(1.50)	Gd.(0.75)	Gd.(1.50)	Gd.(0.75)	Gd.(0.75)	Gd.(0.75)	Gd.(0.75)	Gd.(1.50)	Gd.(0.75)	Fr.(1.50)
Av.(1.00)	Av.(0.50)	Av.(1.00)	Av.(0.50)	Av.(0.50)	Av.(0.50)	Av.(0.50)	Av.(1.00)	Av.(0.50)	Av.(1.00)
Pr.(0.50)	Pr.(0.25)	Pr.(0.50)	Pr.(0.25)	Pr.(0.25)	Pr.(0.25)	Pr.(0.25)	Pr.(0.50)	Pr.(0.25)	Hv.(0.50)
Comment../14									

Organisation/ 7			
SI(1/1)	**EI(2/2)**	**CbI(2/2)**	**SO/TO (2/2)**
Ex. (1.00)	Ex. (2.00)	Ex. (2.00)	Ex. (2.00)
Gd.(0.75)	Gd.(1.50)	Gd.(1.50)	Gd.(1.50)
Av.(0.50)	Av.(1.00)	Av.(1.00)	Av.(1.00)
Pr.(0.25)	Pr.(0.50)	Pr.(0.50)	Pr.(0.50)
Comment………………………./7			

Content/ 6			
TS(2/2)	**SS(2/2)**	**CS(1/1)**	**RtT(1/1)**
Ex. (2.00)	Ex. (2.00)	Ex. (1.00)	Ex. (1.00)
Gd.(1.50)	Gd.(1.50)	Gd.(0.75)	Gd.(0.75)
Av.(1.00)	Av.(1.00)	Av.(0.50)	Av.(0.50)
Pr.(0.50)	Pr.(0.50)	Pr.(0.25)	Pr.(0.25)
Comment……………………......./6			

Final Assessment ....

Appendix 12: Working Portfolio Project

Working Portfolio Project **Student Full Name:……….…………......**

I-Take into consideration the instructions in the column in the table below, then write appropriate responses.

Practice	**Genre** **Instructions**	**Descriptive (D)**	**Narrative (N)**
Writing	Order your **writings** from the best to the poorest one.	D…../D…../D……	N……/N……/N…….
	Why have you elected D……/N…..as the **poorest** ones?	(D…..) is the poorest one because ………………………………… ………………………………… ………………………………… …………………………………..	(N…..) is the poorest one because………..……………… …………………………………. ………………………………… ………………………………….
	Why have you elected D……/N…..as the **best** ones?	(D…..) is the best one because ………………………………… ………………………………… ………………………………......	(N…..) is the best one because ………………………………… ………………………………… ………………………………… ………………………………….
Self-assessment	Order your **self-assessments** from the best to the poorest one	D…../D…../D……	N……/N……/N…….
	Why have you elected D……/N…..as the **poorest** ones?	(D…..) is the poorest one because ………………………………… ………………………………… ………………………………… ………………………………….	(N…..) is the poorest one because………………………… …………………………………. ………………………………… …………………………………
	Why have you elected D……/N…..as the **best** ones?	(D…..) is the best one because ………………………………… ………………………………… ………………………………… ………………………………… …………………………………	(N…..) is the best one because ………………………………… ………………………………… ………………………………… ………………………………… …………………………………
Anonymous Peer	Order the **anonymous peer assessment**you received from	D…../D…../D……	N……/N……/N…….

	the best to the poorest one		
	Why have you elected D……/N…..as the **poorest** ones?	(D…..) is the poorest one because ………………………………… ………………………………… ………………………………… …………………………………	(N…..) is the poorest one because……………………….. ………………………………… …………………………………
	Why have you elected D……/N…..as the **best** ones?	(D…..) is the best one because ………………………………… ………………………………… ………………………………… ………………………………… …………………………………	(N…..) is the best one because ………………………………… ………………………………… ………………………………… ………………………………… …………………………………
Re-writing	Order your **rewritings** from the best to the poorest one (If you had done any. If not, please skip this section)	D…../D…../D……	N……/N……/N…….
	Why have you elected D……/N…..as the **poorest** ones?	(D…..) is the poorest one because ………………………………… ………………………………… ………………………………… ………………………………… …………………………………	(N…..) is the poorest one because ………………………………… ………………………………… ………………………………… …………………………………
	Why have you elected D……/N…..as the **best** ones?	(D…..) is the best one because ………………………………… ………………………………… …………………………………	(N…..) is the best one because ………………………………… ………………………………… …………………………………

II-What do you think about the idea of collecting all your works? (For example: what did you learn from this portfolio? Did you dis/like it? etc.)……………………………………………………………………………………

……………………………………………………………………………………………

……………………………………………………………………………………………

……………………………………………………………………………………………

……………………………………………………………………………………………

……………………………………………………………………………………………

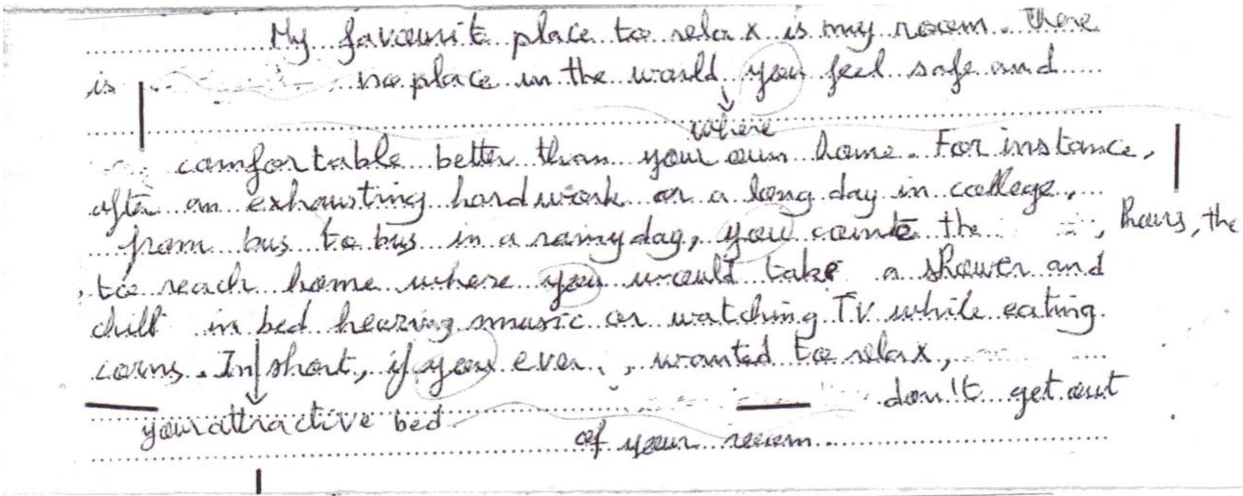
……………………………………………………………………………………………

Thank You Very Much for Your Contribution

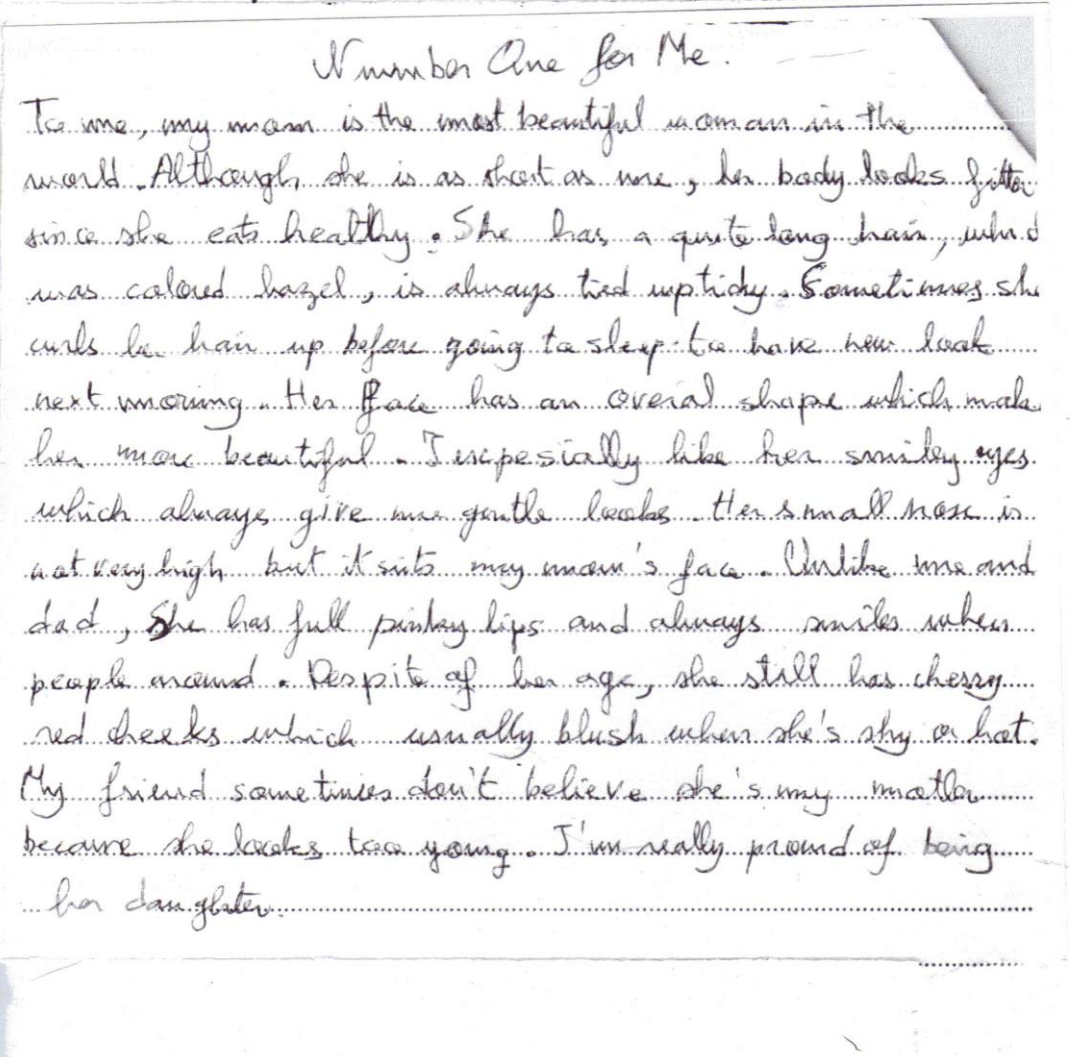
Appendix 13: Participants' Writing Progress in the Intervention Group

Students' Writings in the Beginning and at the End

Student with Poor Level



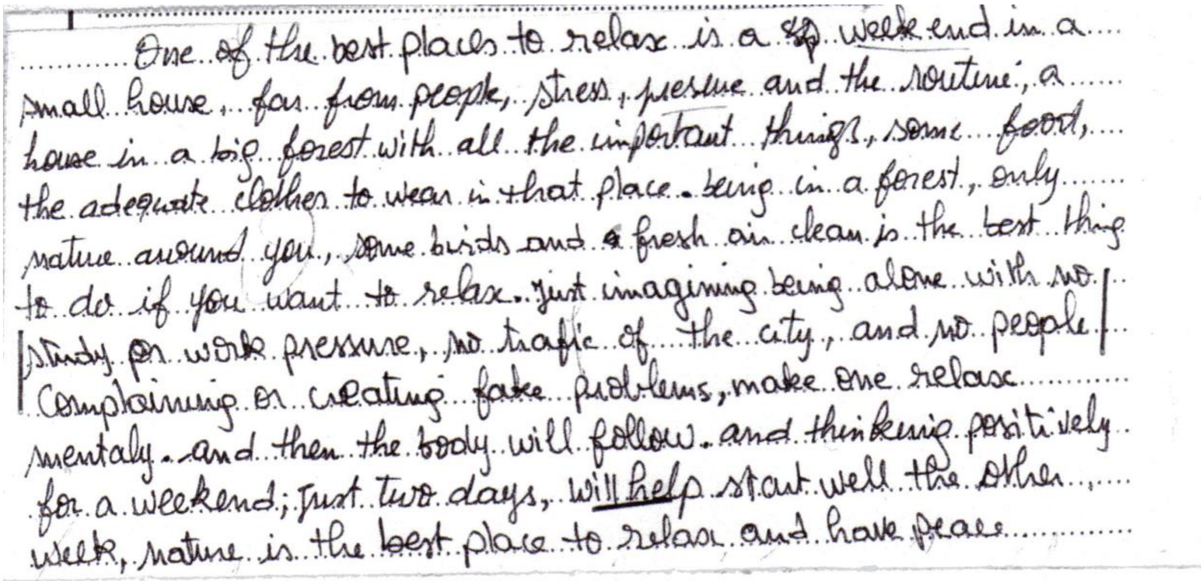
My favourit place to relax is my room. There
is no place in the world you feel safe and
where
comfortable better than your own home. For instance,
after an exhausting hardwork or a long day in college,
from bus to bus in a rainy day, you counte the , hours, the
, to reach home where you would take a shower and
chill in bed hearing music or watching T.V while eating
corns. In short, if you ever, wanted to relax, don't get out
your attractive bed
of your room.



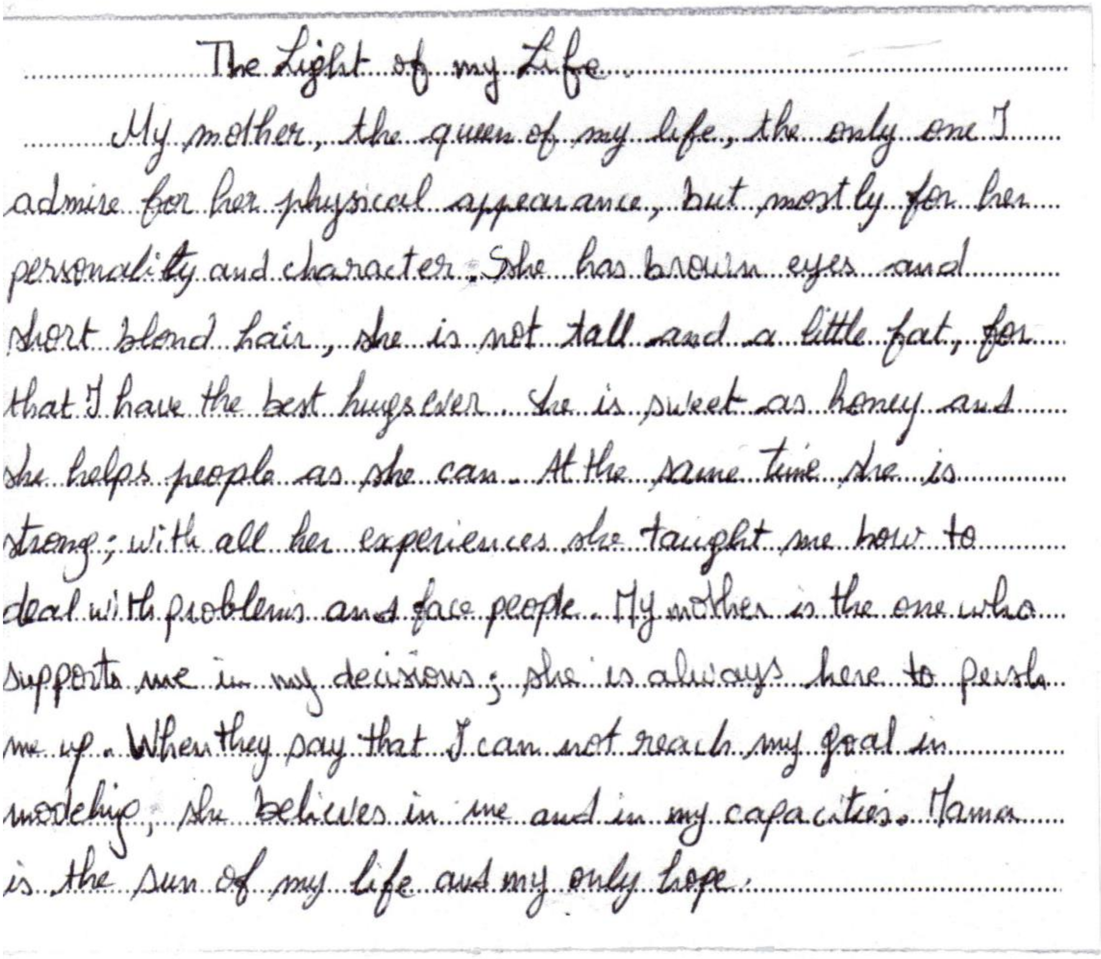
Number One for Me.

To me, my mom is the most beautiful woman in the
world. Although she is as short as me, her body looks fitter
since she eats healthy. She has a quite long hair, which
was colored hazel, is always tied up tidy. Sometimes sh
curls her hair up before going to sleep to have new look
next morning. Her face has an overial shape which mak
her more beautiful. I espesially like her smiley eyes
which always give me gentle looks. Her small nose is
not very high but it suits my mom's face. Unlike me and
dad, She has full pinky lips and always smiles when
people around. Despite of her age, she still has cherry
red cheeks which usually blush when she's shy or hot.
My friend sometimes don't believe she's my mother
because she looks too young. I'm really proud of being
her daughter.

Student with Average Level



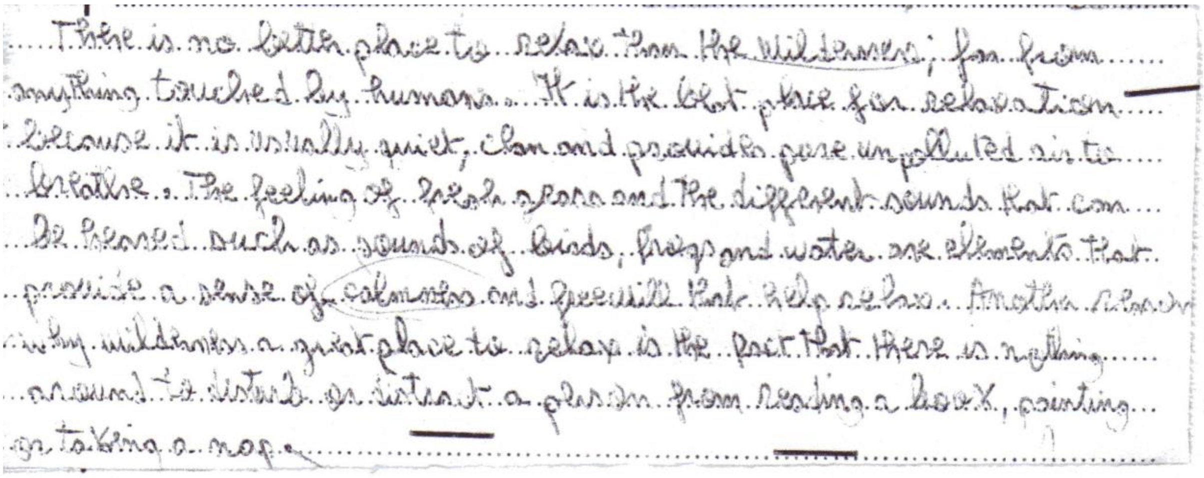
One of the best places to relax is a ~~sf~~ week end in a small house, far from people, stress, presure and the routine; a house in a big forest with all the important things, some food, the adequate clothes to wear in that place. being in a forest, only nature around you, some birds and a fresh air clean is the best thing to do if you want to relax. Just imagining being alone with no study or work pressure, no trafic of the city, and no people complaining or creating fake problems, make one relax mentaly. and then the body will follow. and thinking positively for a weekend; just two days, will help start well the other week, nature is the best place to relax and have peace



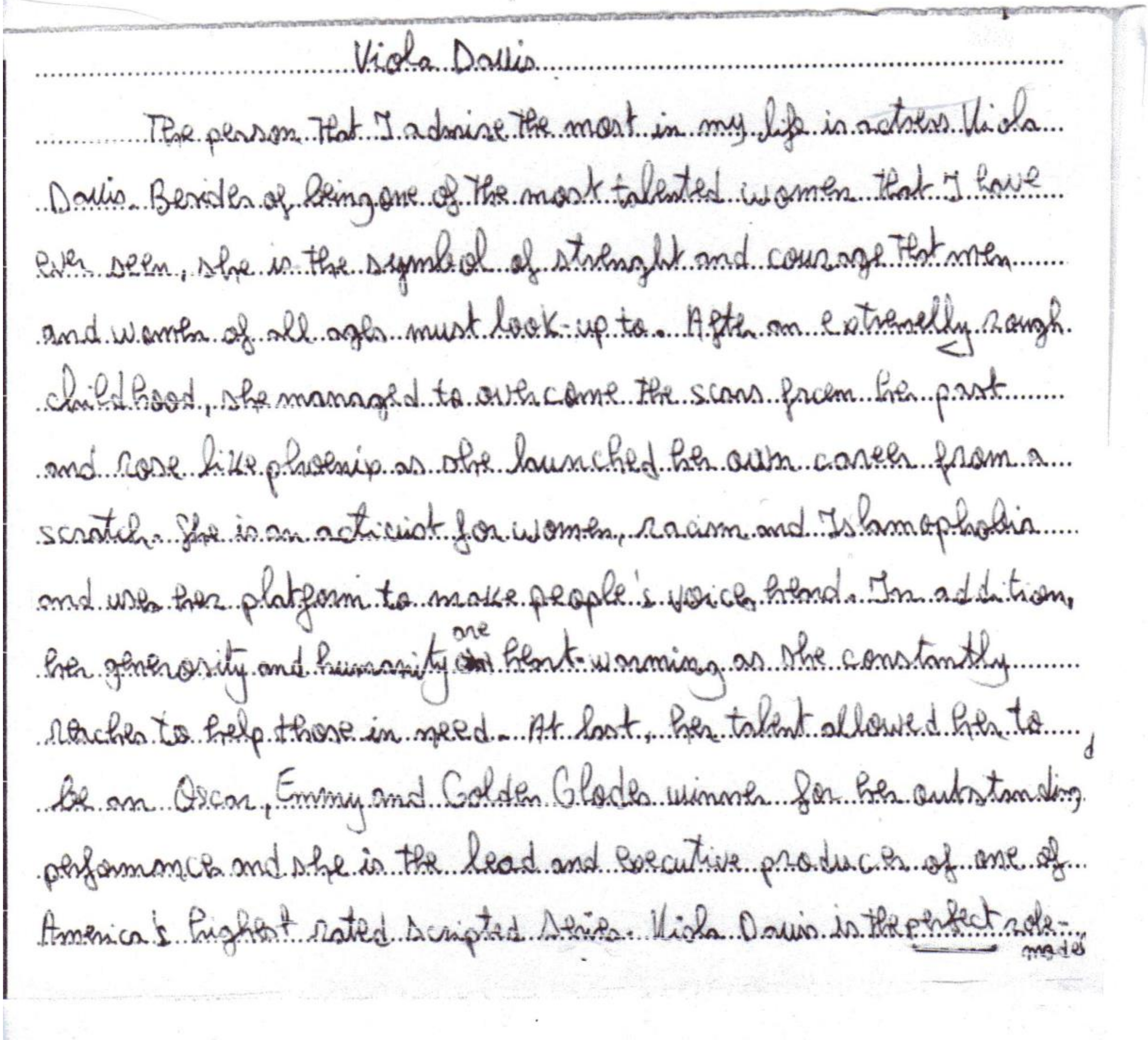
The Light of my Life

My mother, the queen of my life, the only one I admire for her physical appearance, but mostly for her personality and character. She has brown eyes and short blond hair, she is not tall and a little fat, for that I have the best hugs ever. She is sweet as honey and she helps people as she can. At the same time she is strong; with all her experiences she taught me how to deal with problems and face people. My mother is the one who supports me in my decisions; she is always here to push me up. When they say that I can not reach my goal in modeling, she believes in me and in my capacities. Mama is the sun of my life and my only hope.

Student with Good Level



There is no better place to relax than the wilderness; far from anything touched by humans. It is the best place for relaxation because it is usually quiet, clean and provides pure unpolluted air to breathe. The feeling of fresh grass and the different sounds that can be heard such as sounds of birds, frogs and water are elements that provide a sense of calmness and freewill that help relax. Another reason why wilderness a great place to relax is the fact that there is nothing around to disturb or distract a person from reading a book, painting or taking a nap.



Viola Davis

The person that I admire the most in my life is actress Viola Davis. Besides of being one of the most talented women that I have ever seen, she is the symbol of strength and courage that men and women of all ages must look up to. After an extremely rough childhood, she managed to overcome the scars from her past and rose like phoenix as she launched her own career from a scratch. She is an activist for women, racism and Islamophobia and uses her platform to make people's voices heard. In addition, her generosity and humanity ~~an~~ are heart-warming as she constantly reaches to help those in need. At last, her talent allowed her to be an Oscar, Emmy and Golden Globes winner for her outstanding performance and she is the lead and executive producer of one of America's highest rated scripted series. Viola Davis is the perfect role-models

Appendix 14 : Analysis Methods Used

a. Analysis Methods Used in the First Stage of Data Analysis Process

Figure/Table Number	**Analysis Method Used**	**Characteristics**
Figure 3.4	Content Analysis	Categories/ Manifest/ Quantitative/ Inductive &Deductive
Figure 3.6	Content Analysis	Categories/ Manifest &Latent/ Quantitative/ Inductive
Figure 3.7	Content Analysis	Categories/ Manifest/ Quantitative/ Inductive
Figure 3.8	Content Analysis	Categories/ Manifest & Latent/ Quantitative/ Inductive
Figure 3.9	Content Analysis	Categories/ Manifest/ Quantitative /Inductive
Figure 3.10	Content Analysis	Categories/ Manifest/ Quantitative /Inductive
Figure 3.19	Content Analysis	Categories/Manifest/ Quantitative/Inductive
Figure 3.22	Thematic Analysis	Themes/ Latent/ Quantitative/ Inductive
Figure 3.23	Content Analysis	Categories/ Manifest/ Quantitative/ Inductive &Deductive
Figure 3.25	Content Analysis	Categories/ Manifest/ Quantitative/ Inductive
Figure 3.26	Blended Methods	Themes & Categories/ Manifest & Latent/ Quantitative/ Inductive &Deductive
Figure 3.28	Content Analysis	Categories/ Manifest/ Quantitative/ Inductive & Deductive
Figure 3.29	Content Analysis	Categories/ Manifest/ Quantitative/ Inductive & Deductive
Figure 3.31	Content Analysis	Categories/ Manifest &Latent/ Quantitative/ Inductive & Deductive
Figure 3.32	Content Analysis	Categories/ Manifest & Latent/ Quantitative/ Inductive
Figure 3.34	Content Analysis	Categories/ Manifest/ Quantitative/ Inductive
Figure 3.36	Content Analysis	Categories/ Manifest/ Quantitative/ Inductive
Figure 3.37	Content Analysis	Categories/ Manifest/ Quantitative/ Inductive
Figure 3.39	Content Analysis	Categories/ Manifest/ Quantitative/ Inductive

Figure 3.40	Content Analysis	Categories/ Manifest/ Quantitative/ Inductive &Deductive
Figure 3.42	Content analysis	Categories/ Manifest/ Quantitative/ Inductive
Table 3.1	Thematic Analysis	Categories/ Manifest/ Quantitative/ Inductive & Deductive

b. Analysis Methods Used in the Third Stage of Data Analysis Process

Figure/Table Number	**Analysis Method Used**	**Characteristics**
Figure 3.45	Content Analysis	Categories/ Manifest/ Quantitative/ Inductive
Figure 3.50	Content Analysis	Categories/ Manifest/ Quantitative/ Inductive & Deductive
Figure 3.51	Content Analysis	Categories/ Manifest/ Quantitative/ Inductive & Deductive
Figure 3.52	Content Analysis	Categories/ Manifest & Latent/ Quantitative/ Inductive & Deductive
Figure 3.53	Content Analysis	Categories/ Manifest/ Quantitative/ Inductive & Deductive
Figure 3.54	Content Analysis	Categories/ Manifest & Latent/ Quantitative/ Inductive
Figure 3.55	Content Analysis	Categories/ Manifest & Latent/ Quantitative/Inductive & Deductive
Figure 3.56	Content Analysis	Categories/ Manifest/Quantitative/ Inductive
Figure 3.57	Content Analysis	Content Analsysis/ Manifest/ Qunatitative/ Inductive
Figure 3.58	Content Analysis	Categories/ Manifest & Latent/ Quantitative/ Inductive
Figure 3.59	Content Analysis	Categories/ Manifest/ Quantitative/ Inductive & Deductive
Figure 3.60	Content Analysis	Categories/ Manifest & Latent/ Quantitative/ Inductive & Deductive
Figure 3.62	Content Analysis	Categories/ Manifest/ Quantitative/ Inductive

		&Deductive
Figure 3.63	Content Analysis	Categories/ Manifest/ Quantitative/ Inductive & Deductive
Figure 3.64	Content Analysis	Categories/ Manifest/ Quantitative/ Inductive
Figure 3.67	Content Analysis	Categories/ Manifest/ Quantitative/Inductive
Table 3.30	Blended Method	Themes & Categories/ Manifest/ Quantitative/ Inductive & Deductive
Table 3.31	Content Analysis	Categories/ Manifest/ Quantitative/ Inductive & Deductive
Table 3.32	Content Analysis	Categories/ Manifest/ Quantitative & Qualitative/ Inductive & Deductive
Table 3.33	Content Analysis	Ctegories/Manifest & Latent/ Quntitative & Qualitative/ Inductive & Deductive
Table 3.34	Content Analysis	Categories/ Manifest & Latent/ quantitative and qualitative/ Inductive & Deductive
Table 3.36	Content Analysis	Categories/ Manifest/ Quantitative/ Inductive & Deductive
Table 3.37	Content Analysis	Categories/ Manifest/ Quantitative/ Inductive
Table 3.39	Blended Method	Categories & Themes/ Manifest & Latent/ Quantitative/ Inductive & Deductive
Table 3.40	Content Analysis	Categories/ Manifest & Latent/ Quantitative Qualitative/ Inductive & Deductive
Table 3.41	Blended Method	Categories & Themes/ Manifest & Latent/ Quantitative & Qualitative/ Inductive & Deductive
Table 3.42	Blended Method	Categories & Themes/ Manifest & Latent/ Quantitative/ Inductive & deductive
Table 3.43	Blended Method	Categories & Themes/ Manifest & Latent/ Quantitative & Qualitative/ Inductive & Deductive
Table 3.44	Content Analysis	Categories/ Manifest/ Quantitative/ Inductive
Table 3.46	Content Analysis	Categories/ Manifest/ Quantitative/ Inductive