ACHIEVING OBJECTIVITY IN ASSESSING STUDENTS’ ESSAYS: TOWARDS AN AUTOMATIC WRITING ASSESSMENT

The case of the third year English LMD students at the University of Sidi-bel-Abbes

Thesis Submitted to the Department of English in Candidacy for the Degree of Doctorate in Didactics of Specialism

Presented by: Mr. Abdellah BARAKA
Supervised by: Prof. Mohamed MELOUK

Board of Examiners:

- Prof. Zouaoui MERBOUH (Professor) President (University of Sidi-bel-Abbes)
- Prof. Mohamed MELOUK (Professor) Supervisor (University of Sidi-bel-Abbes)
- Prof. Belabbes OUERRAD (Professor) Internal Examiner (University of Sidi-bel-Abbes)
- Dr. Habib YAHIAOUI (MCA) External Examiner (University of Mascara)
- Dr. Mohamed GRAZIB (MCA) External Examiner (University of Saida)
- Dr. Azzeddine BOUHASSOUN (MCA) External Examiner (University of Ain Temouchent)

Academic Year: 2015-2016
Dedications

To my family,

“The only rock I know that stays steady, the only institution I know that works”

Lee Iacocca

To my teachers,

“I can no other answer make, but, thanks, and thanks.”

William Shakespeare

To my friends,

But if the while I think on thee, dear friend, all losses are restored and sorrows end.

William Shakespeare
Acknowledgements

First and foremost, I would like to thank Allah, the almighty, for granting me guidance, strength and courage to complete this thesis. Second, because *feeling gratitude and not expressing it is like wrapping a present and not giving it* (William Arthur), I would like to articulate my truthful gratitude to my supervisor Prof. Mohamed MELOUK for the incessant assistance of my doctorate revise and related research, for his tolerance, patience, inspiration, motivation and immense knowledge. He was available during every fraction of the time when such humble research was in progression. I could never dare imagine having a better supervisor and counselor for my doctorate writing. As, I would like to thank the rest of my thesis committee: Prof. Zouaoui MERBOUH, Prof. Belabbes OUERRAD, Dr. Habib YAHIAOUI, Dr. Azzeddine BOUHASSOUN and Dr. Mohamed GRAZIB not only for their perceptive remarks and support, but also for their firm inquiries which inspire me to broaden my research towards various standpoints. I am also deeply indebted to Prof. Andrew BREEZE and Prof. Ruth BREEZE, from the University of Navarra, Spain for providing me with enlightenment, invaluable support, brilliant guidance, enthusiasm and encouragement throughout the research. As I owe sincere and earnest thankfulness and appreciations to Prof. Zouaoui MERBOUH, who offered me a chance to join his team of research. Without his precious support it would not be possible to conduct this investigation. I would like also to express my deep gratitude and indebtedness to my students and colleagues at the Universities of Mascara and Sidi-Bel-Abbes for the interesting discussions, scientific collaborations and for all the fun we have had during the last five years. Last but not the least; I would like to thank my mother, my wife, and my children for enlightening me the first glance of research.
Abstract

Shaping the most effective approach for the assessment of the writing abilities has become a challenge for both teachers and researchers. Investigations about an effective and reliable assessment to the learners’ writing have resulted in the requirement of an atomistic assessing organism, which may make the assessing process more objective and reliable to interfere so as to enhance the learners’ writing skills. To embark upon this study, the researcher was undeniably spurred by the ineffectiveness of the instructor’s assessment, which is often questioned by both parties. The learners often perceive that their advancement is slow with the same hindrances recurring, while the teachers become bothered at noticing the same unsuitable features occurring yet again. The present study aims at enhancing the teacher’s assessment reliability to the written language through the generation of an analytic assessing procedure. Such a procedure may reflect every detail about the writing performance of the learner. The study also points out to the possibility of involving computer technology to enhance the reliability of such kind of analytic assessment process. In order to undertake such a task a range of data and different approaches have been used. The data used in this work come, first, from a corpus of linguistics examination papers of the second and third year English LMD students at the University of Sidi-bel-Abbes and second from the linguistics teachers’ interview about the process of writing assessment. The methodology followed in this work is a qualitative comparative analysis based on a psychometric /atomistic approach. The present study has revealed that most teachers are assessing their learners’ writing holistically therefore no feedback is designed and no improvement is taking place, the fact that makes such kind of assessment never achieve reliability. On the other hand, the use of the analytic procedure associated with computer technology to assess writing may enhance and increase the reliability, the validity and the effectiveness of the teachers’ assessment as well as the learners’ writing performance.
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AEA</td>
<td>Automatic essay assessment</td>
</tr>
<tr>
<td>AES</td>
<td>Automated Essay Scoring</td>
</tr>
<tr>
<td>AI</td>
<td>Artificial Intelligence</td>
</tr>
<tr>
<td>BETSY</td>
<td>Bayesian Essay Test Scoring System</td>
</tr>
<tr>
<td>EFL</td>
<td>English as Foreign Language</td>
</tr>
<tr>
<td>E-rater</td>
<td>Electronic Raster</td>
</tr>
<tr>
<td>ESL</td>
<td>English as Foreign Language</td>
</tr>
<tr>
<td>FP</td>
<td>Feedback about the Processing of the task</td>
</tr>
<tr>
<td>FR</td>
<td>Feedback about self Regulation</td>
</tr>
<tr>
<td>FS</td>
<td>Feedback about the Self</td>
</tr>
<tr>
<td>FT</td>
<td>Feedback about the Task</td>
</tr>
<tr>
<td>GMAT</td>
<td>Graduate Management Admissions Test</td>
</tr>
<tr>
<td>IEA</td>
<td>Intelligent Essay Assessor</td>
</tr>
<tr>
<td>LMD</td>
<td>Licence, Master and Doctorate</td>
</tr>
<tr>
<td>NI</td>
<td>Natural Intelligence</td>
</tr>
<tr>
<td>NLP</td>
<td>Natural Language Processing</td>
</tr>
<tr>
<td>PEG</td>
<td>Project Essay Grader</td>
</tr>
<tr>
<td>TOEFL</td>
<td>Test Of English as a Foreign Language</td>
</tr>
</tbody>
</table>
List of tables

Chapter One

Table 1.1. Scoring Model Adapted from Oregon Department of Education 2005

Table 1.2. Writing performance appreciations (adapted from Moskal 2000)

Table 1.3. Third year teacher adapted Linguistics syllabus (1st Semester)

Table 1.4. Third year teacher adapted Linguistics syllabus (2nd Semester)

Table 1.5. Second year teacher adapted Linguistics syllabus (1st Semester)

Table 1.6. Second year teacher adapted Linguistics syllabus (2nd Semester)

Table 1.7. The researcher’s assessments

Table 1.8. Teachers’ assessing time

Table 1.9. The researcher’s threefold assessment

Table 1.10. The teachers’ assessment through the analytic method

Table 1.11. The researcher methodology

Chapter three

Table 3.1. Second year written test results (Passing marks /Group A)

Table 3.2. Second year written test results (Failing marks /Group A)

Table 3.3. Second year written test results (Passing marks / Group B)

Table 3.4. Second year written test results (Failing marks /Group B)

Table 3.5. Group “A” success rates
Table 3.6. The assessment unreliability

Table 3.7. The selected essays samples

Table 3.8. The selected students as assessors

Table 3.9. The students’ assessments to their peers’ writings

Table 3.10. The whole assessment procedure

Table 3.11. The group ‘A’ marks after the three holistic assessment attempts

Table 3.12. The group ‘B’ marks after the three holistic assessment attempts

Table 3.13. Group “A” passing and failing rates

Table 3.14. Group “B” passing and failing rates

Table 3.15. The unreliability of the success rates of the whole population

Table 3.16. The selected essays analysis samples

Table 3.17. The teachers’ assessment to the selected written essays

Table 3.18. Teacher 1 rates to each an assessment scale

Table 3.19. Teacher 2 rates to each an assessment scale

Table 3.20. Teacher 3 rates to each an assessment scale

Table 3.21. Teachers’ assessed written components

Chapter Four

Table 4.1. The teacher expected content organisation (1)

Table 4.2. The teacher expected content organisation (2)
Table 4.3. The teacher expected content organisation (3) 197

Table 4.4. The teacher expected content organisation (4) 198

Table 4.5. The teacher expected content organisation (5) 198

Table 4.6. The teacher’s assessing grid example 201

Table 4.7. The researcher’s adapted assessing grid 220
List of Figures

Chapter One

Figure 1.1. Writing Class Pattern (Adapted from Raimes, 1983) 28

Figure 1.2. Analytic versus holistic methods, contrastive graph 36

Figure 1.3. Sampling Frame 51

Figure 1.4. The students assessing procedure 53

Figure 1.5. The researcher’s adapted assessing procedure 59

Chapter Two

Figure 2.1. Race’s ripple on a pond modal 103

Chapter Three

Figure 3.1. Sampling frame 137

Chapter Four

Figure 4.1. The researcher’s tailored analytic assessing process 192

Figure 4.2. The 3rd year written test’s answer sample 193

Figure 4.3. Sample Feedback of Intelligent Essay Assessor 207

Figure 4.4. Hattie & Timperley model of an effective feedback 212

Figure 4.5. The assessing procedure 219
List of graphs

Chapter one

Graph 1.1. Madsen’s Analytic Scoring Method

Chapter three

Graph 3.1. The group (A) success and failure rates
Graph 3.2. The total points allocated to group (A)
Graph 3.3. The marks unreliability in group (A)
Graph 3.4. The degree of reliability
Graph 3.5. The final assessment provided by the whole participants
Graph 3.6. The unreliability of the passing and failing marks
Graph 3.7. The unreliability of the total points
Graph 3.8. The unreliability of the assessing procedures
Graph 3.9. The marks’ unreliability
Graph 3.10. The marks’ unreliability (Group B)
Graph 3.11. The assessing procedure unreliability
Graph 3.12. The essay samples marks unreliability
Graph 3.13. The three assessments results
Graph 3.14. The first assessment results
Graph 3.15. The degree of unreliability
Graph 3.16. The unreliability of the multiple assessing procedures (Essay A)

Graph 3.17. The unreliability of the multiple assessing procedures (Essay B)

Graph 3.18. Essay “A” assessment unreliability

Graph 3.19. Essay “B” assessment unreliability

Graph 3.20. Essay “A” assessment unreliability percentage

Graph 3.21. Essay “B” assessment unreliability percentage
# Table of Content

Dedication I  
Acknowledgement II  
Abstract III  
List of Abbreviations IV  
List of Tables V  
List of Figures VIII  
List of Graphs IX  
Table of Content XI  
General Introduction 01

## Chapter One The Detection of The Complication

1.1. Introduction 10  
1.2. The Implementation of the LMD System 11  
   1.2.1. LMD Features and Objectives 13  
   1.2.2. English Language Teaching Within the LMD 14  
1.3. Competency Based Approach 15  
   1.3.1. Characteristics of CBA 16  
   1.3.2. The Implementation of CBA in Algeria 17  
1.4. Teaching Writing through CBA 19  
   1.4.1. The Writing Deficits 20  
   1.4.2. Writing Feedback 24  
   1.4.3. Responding to Learners’ Writing 27  
1.5. Assessing Writing 30  
   1.5.1. Analytic Assessing Method 30  
   1.5.2. Holistic Assessing Method 35  
1.6. The Research Type 37
1.6.1. The Research Purposes
1.6.2. The Research Problematic
1.6.3. The Research Questions
1.6.4. The Research Hypotheses

1.7. The Research Population
1.7.1. Main Population / Third Year Students
1.7.2. The Main Population Linguistics Syllabus
1.7.3. Secondary Population / Second Year Students
1.7.4. The Secondary Population Linguistics Syllabus

1.8. Data Collection Devices
1.9. Secondary Population’s Written Test
1.9.1. The Students’ Contributions to the Assessment Procedure
1.9.2. The Selection of the Written-Papers
1.9.3. The Selection of Students as Assessors
1.9.4. The Assessing Procedure:
   1.9.4.1. The First Procedure: Self-Assessment
   1.9.4.2. The Second Procedure: Assessment Exchange
   1.9.4.3. The Third Procedure: Multiple-Assessment

1.10. Primary Population Written Test
1.10.1. The First Tentative Assessment: Holistic Assessment
   1.10.1.1. The Researcher’s Assessments
   1.10.1.2. The Teachers’ Contributions: Multiple Assessments
1.10.2. The Second Tentative Assessment: Analytic Assessment
   1.10.2.1. The Researcher’s Assessment: The Threefold Assessment
   1.10.2.2. The Teachers’ Assessments

1.11. The Teachers’ Interview
Chapter Two Enhancing the Essay Assessment Reliability

2.1. Introduction

2.2. Testing the Written Production
   2.2.1. Objective Testing
   2.2.2. Subjective Testing “The Essay”

2.3. Subjective Testing Specificities
   2.3.1. Scoring Unreliability
   2.3.2. Validity Deficiency
   2.3.3. Interpretation Difficulty
   2.3.4. Preparation Effortlessness
   2.3.5. Administration Simplicity
   2.3.6. Scoring Difficulty

2.4. The Essay Assessment Approaches

2.5. The Holistic Approach
   2.5.1. The Holistic Approach Deficits
   2.5.2. The Holistic Assessment New Forms
      2.5.2.1. The Administration of Short Essays
      2.5.2.2. The Multiple Assessing Method
      2.5.2.3. The Error- Count Method

2.6. The Analytical Assessing Method
   2.6.1. The Nature of the Analytic Assessing Method
   2.6.2. The Analytic Assessing Procedure
   2.6.3. The Analytic Assessing Method Deficiency
   2.6.4. The Analytic Assessing Method Advantages
2.6.4.1. Scores Reliability

2.6.4.2. Security

2.6.4.3. Feedback

2.7. Feedback: An Overview

2.8. The Feedback Effectiveness and Ineffectiveness

2.9. The Feedback Levels

2.9.1. Feedback about the Task

2.9.2. Feedback about the Process

2.9.3. Feedback about Self-Regulation

2.9.4. Feedback about the Self

2.10. Writing Feedback

2.11. The Writing Feedback Forms

2.11.1. Direct Corrective Feedback

2.11.2. Indirect Corrective Feedback

2.11.3. Metalinguistic Corrective Feedback

2.11.4. Focused Versus Unfocused Feedback

2.12. The Teacher’s Roles in Delivering Feedback

2.13. Conclusion

Chapter Three Scrutinizing the Findings

3.1. Introduction

3.2. The Second Year Written Test Assessment

3.2.1. The Researcher’s Assessing Procedure

3.2.1.1. The Success and Failure Rates Reliability

3.2.1.2. The Researcher’s Assessment Unreliability

3.2.2. Students' Assessing Procedure
3.2.2.1. The Sampling Frame
3.2.2.2. The Students’ Assessment Contribution

3.3. The Third Year Written Test Assessment
3.3.1. The Impressionistic Assessing Procedure
3.3.2. The Researcher’s Assessment
   3.3.2.1. The Unreliability of the Passing / Failing Marks: Group A
   3.3.2.2. Unreliability of the Assessing Procedure:
3.3.3. Essays’ Samples Analysis
3.4. Teachers’ Contributions (Multiple Scorings)
3.5. The Teachers’ Interview
3.6. The Results’ Interpretation
3.7. Conclusion

Chapter Four Pedagogical Implications

4.1. Introduction
4.2. The Formal Assessing Procedure
   4.2.1. The Selection of the Assessing Components
   4.2.2. The Specification of the Criteria of Assessment
      4.2.2.1. The Relevance
      4.2.2.2. Paragraphing
      4.2.2.3. Written Expression
      4.2.2.4. Organisation and Coherence
      4.2.2.5. Grammar and Mechanics
   4.2.3. The Architecture of an Analytic Assessing Scheme
4.3. The Macrostructural Assessing Procedure
4.4. The Microstructural Assessing Procedure
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.4.1. Assessing the Paragraphing of Ideas</td>
<td>195</td>
</tr>
<tr>
<td>4.4.2. Assessing the Organisation and Coherence of Ideas:</td>
<td>199</td>
</tr>
<tr>
<td>4.4.3. Assessing Language Form</td>
<td>200</td>
</tr>
<tr>
<td>4.4.4. Assessing the Style</td>
<td>200</td>
</tr>
<tr>
<td>4.5. The Analytic Assessing Limitations</td>
<td>202</td>
</tr>
<tr>
<td>4.6. Automatic Essay Assessment</td>
<td>203</td>
</tr>
<tr>
<td>4.6.1. Project Essay Grader (PEG)</td>
<td>205</td>
</tr>
<tr>
<td>4.6.2. Intelligent Essay Assessor</td>
<td>206</td>
</tr>
<tr>
<td>4.6.3. E-Rater</td>
<td>207</td>
</tr>
<tr>
<td>4.6.4. IntelliMetric</td>
<td>208</td>
</tr>
<tr>
<td>4.6.5. Bayesian Essay Test Scoring System</td>
<td>208</td>
</tr>
<tr>
<td>4.7. Designing an Effective Feedback</td>
<td>210</td>
</tr>
<tr>
<td>4.8. Reshaping the Writing Feedback</td>
<td>212</td>
</tr>
<tr>
<td>4.8.1. Where Am I Going?</td>
<td>213</td>
</tr>
<tr>
<td>4.8.2. How Am I Going?</td>
<td>213</td>
</tr>
<tr>
<td>4.8.3. Where to Next?</td>
<td>214</td>
</tr>
<tr>
<td>4.9. Enhancing the Effectiveness of the Feedback.</td>
<td>214</td>
</tr>
<tr>
<td>4.9.1. Reflecting Upon the Feedback Process</td>
<td>215</td>
</tr>
<tr>
<td>4.9.2. Considering Students’ Needs and Suggestions</td>
<td>215</td>
</tr>
<tr>
<td>4.9.3. Varying Feedback Brands</td>
<td>216</td>
</tr>
<tr>
<td>4.9.4. Feeding Up</td>
<td>217</td>
</tr>
<tr>
<td>4.9.5. Equilibrating Positive and Negative Feedback</td>
<td>217</td>
</tr>
<tr>
<td>4.9.6. Feeding Forward</td>
<td>218</td>
</tr>
<tr>
<td>4.9.7. Incorporating Remedial Task</td>
<td>218</td>
</tr>
<tr>
<td>4.10. Planning a Corrective Conference</td>
<td>219</td>
</tr>
<tr>
<td>4.11. Managing the Corrective Conference</td>
<td>220</td>
</tr>
</tbody>
</table>
4.11.1. The Review of the Essay Question
4.11.2. Teacher’s Impressions
4.11.3. Collective Correction
4.11.4. Individual Correction
4.11.5. Remedial Work

4.12. Conclusion

General Conclusion
References

Appendices

<table>
<thead>
<tr>
<th>Appendix</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appendix 1</td>
<td>Second Year Written Test</td>
<td>241</td>
</tr>
<tr>
<td>Appendix 2</td>
<td>Third Year Written Test</td>
<td>243</td>
</tr>
<tr>
<td>Appendix 3</td>
<td>Second Year Written Test Answer Sample</td>
<td>244</td>
</tr>
<tr>
<td>Appendix 4</td>
<td>Third Year Written Test Answer Sample</td>
<td>246</td>
</tr>
<tr>
<td>Appendix 5</td>
<td>Selected Cloze Test “A”</td>
<td>247</td>
</tr>
<tr>
<td>Appendix 6</td>
<td>Selected Cloze Test “B”</td>
<td>249</td>
</tr>
<tr>
<td>Appendix 7</td>
<td>Selected Written Essay “A”</td>
<td>251</td>
</tr>
<tr>
<td>Appendix 8</td>
<td>Selected Written Essay “B”</td>
<td>252</td>
</tr>
<tr>
<td>Appendix 9</td>
<td>The Analytic Assessing Formula</td>
<td>253</td>
</tr>
<tr>
<td>Appendix 10</td>
<td>The Analytic Scoring Scheme</td>
<td>254</td>
</tr>
<tr>
<td>Appendix 11</td>
<td>The Analytical Assessing Grid</td>
<td>255</td>
</tr>
<tr>
<td>Appendix 12</td>
<td>The Essays Samples Analytical Assessment</td>
<td>256</td>
</tr>
<tr>
<td>Appendix 13</td>
<td>The Teachers’ Interview</td>
<td>257</td>
</tr>
</tbody>
</table>

Glossary
General Introduction
General introduction

As universally conceived, language is the exclusive human faculty. It is that systematic means of interaction that differentiates humans from all other living creatures. Language remains potentially a communicative medium capable of expressing ideas and concepts as well as moods, feelings, attitudes and culture. Today, English has become the most useful and practical language for the global contact. English has acquired the status of a universal language dominating numerous fields. It is the leading language on the Internet, representing over 80% of data and resources. Moreover, the ability to articulate in English makes travelling more serenely and more secure. There is always someone interacting or understanding the language of Shakespeare. English has over time replaced French as the lingua franca of the international diplomacy since the Second World War, although its international importance has never ceased to be represented since the Treaty of Versailles 1919. Indeed, today at least one out of four people have commands over some English basic skills, and one out of two people speak English. Therefore, attaining English proficiency has become the ultimate objective of millions of language learners. They are taking part in countless learning programs all over the world.

Yet, learning English is not always a trouble-free mission mainly in our Algerian socio-cultural context. Researchers are working hard to enhance the situation and make English learning more than just simple. Teaching such hegemonic language has become a prime challenge for teachers who are supposed to teach, assess, score, provide feedback then decide next steps in learning. Acquiring a language, as exposed by a lot of theorists, is a question of going through four basic skills: listening and reading as receptive or input skills and speaking and writing as productive or output skills.
The present enquiry entitled “Achieving objectivity in Assessing the Student’s Essays: is an attempt towards the Automatic Assessment of the Writing. Moreover, it is meant to elucidate a slight issue in the learning of the writing skills namely the assessment reliability. The spur behind such extravagant wish to accomplish this modest work is certainly the sharp deficits in the writing skills of our EFL learners, in particular the hindrances that impede our learners from enhancing their writing performances. The assumption is that such deficits are due to the ineffectiveness of the writing teacher’s feedback. Such feedback ineffectiveness in its turn is unquestionably due to the unreliability and of the assessing process. Therefore, the present research tries to answer three significant questions:

- Is the teacher’s assessment to the EFL learners’ essays reliable enough to achieve effectiveness?
- How can an objective and reliable assessment to writing be achieved?
- Can the computer be an alternative and innovative procedure to enhance the reliability of the essay assessment?

The study aims at finding a solution to the problem of assessing the written language by providing hypotheses that can link the reliability of assessing with the learners’ writing enhancement. The hypotheses call for the design of an analytic assessing system which can control the procedure. Besides, the hypotheses advocate the use of the machine to achieve a better objectivity and reliability of the essay assessment. The proposed hypotheses are attempts to make the assessing of an essay more objective, formal, effective and reliable. Such an assessment may increase the learners’ consciousness over writing complications and provide them with positive feedback over the quality of their writing. Thus, and for triangulation sake, the researcher hypothesizes that:
The teacher’s assessment to his EFL learners’ writing is not reliable and objective so as to achieve effectiveness.

To achieve reliability and objectivity in assessing the EFL learners’ writing, the teacher may map an analytic assessing formula.

Such a designed analytic assessing formula may function more effectively if transformed into a computer programme or software.

To test the validity of the above hypotheses, a number of tools have been used to collect varied data namely two written tests administrated to two different populations at the University of Sidi-bel-Abbes and a teachers’ interview. However, it is worth mentioning that the target population consists of the third year EFL learners, as the experimental population of the study, and the second year EFL learners, as the control population.

Obviously, there are limitless studies which scrutinize formative assessment along with the teacher’s feedback to the EFL learners' writing. Furthermore, countless previous studies have examined different issues about assessment such as its kinds, focus, amount, source, and effectiveness. Yet, the findings related to the problem under investigation are far from being conclusive and decisive. This is due to the complexity of the notion of assessment.

Consequently, the present study brings a humble contribution to the body of research of assessing effectively and objectively the EFL learners' written production. The core of probing into the problem is to investigate the kind of assessment provided by EFL writing teachers, at the University of Sidi-bel-Abbes along with their attitudes and preferences concerning assessment in general. The aim, here, is to develop a conceptual understanding of assessment issues, and highlight the necessity of an effective and reliable assessment for both teachers and learners to enhance teaching and learning of the writing skills in particular. It is
hoped that by conducting such a study, teachers will provide an effective assessment that learners can digest and act on.

Undoubtedly, writing is the final process of each well determined unit in teaching. Learners are required to re-invest the thematic and language element acquired throughout this teaching part, by foregrounding a particular function like advising, comparing, informing, etc. Writing well is one of the most vital talents attempted in learning English as a foreign language. It's not restricted only in the sphere of communication, but it also assists students to develop, order, and unify their assumptions and thoughts since learning to write is tantamount to learning to think. At the same time, it is considered the most multifaceted skill to master since it engages wide ideas in which alterations and thoughts are transformed into written message, writing is not an inborn ability, but rather acquired and practised. Thus, many mistakes are made by foreign language learners in their way to acquire the writing skills, those mistakes determine their deficits in writing as well as evidences of how language learning should be done. Writing makes thinking visible and promotes the ability to pose worth with questions.

Most importantly; writing helps the teacher design an effective feedback. Such a feedback should provide a kind of a channel for the teachers to communicate constructively with students and help them develop as writers. It is obvious enough that the feedback is a sort of facilitator or a vehicle through which the teacher may set up the whole process of teaching. The ideal feedback as described by a lot of theorists should include three responses: feed up, feedback, and feed forward. Simply put, the feedback should provide the student with information concerning, the learning goals, student performance, and ways of improvement. Feedback is not only for learners to close the gap between the actual and the desired performance. It is also crucial for teachers to assess, adjust, and enhance their teaching practices.
Yet, and despite the fact that a large number of research studies has investigated the writing feedback, feedback is still an “enigma” in the field of education as it is miscommunicated and misunderstood among teachers and learners. The problem seems to exist not only in the process of providing feedback, but also in the learners’ attitude towards such a feedback. It is recognised that teachers endow students with mostly written feedback and even oral feedback, but they are indecisive about its functionality or usefulness, and unaware about why it does not work after all. Even worse, the most qualified teachers are still puzzled and perplexed regarding methods of responding competently, professionally, effectively, proficiently, and reasonably to learners' written work.

However and before deciding upon which kind of feedback to be provided, the teacher is to go through a valid testing and a reliable assessment. It is argued that language assessment may be used as (1) a source of information for making decisions within the context of educational programmes; and (2) as an indicator of the learners’ abilities or attributes. In other words, the major uses of tests’ assessment, in educational settings, are related to evaluation or making decisions about either the learners or the syllabuses. Language assessment would provide teachers with reliable and valid measurements for a variety of purposes. They can be used to motivate and guide learners to learn; and because learners are asked to prepare the materials that have been taught, it is the teacher’s responsibility to construct tests that assess important course objectives.

The reliability of the assessing procedure is the central point of the issue under investigation. Assessing the written language seems to be of a paramount significance in the process of acquiring the writing aptitudes since an effective assessment may provide the learner not only with an opportunity to identify and recognise his lacunae but also with sufficient tools to defeat such kind of lacunae. In assessing the written language, the teacher is faced with two recognisable procedures: holistic and analytic assessment. These two
theoretical concepts are distinguished on the basis of the manner the teacher undertakes a
correction session. In testing written language, the learners often deal with two kinds of tests:
the objective test such as the cloze test which has only one correct answer per item. It is easily
scored and assessed since teacher does not waste time to identify his learners’ weaknesses.
Conversely; a free-response test or subjective test requires assessment by making extra efforts
to find out deficiencies. Although teachers are showing an increasing concern in the way they
could provide remedy for their students’ weak performance in writing, they still lack the most
important basis of a formal and effective assessment. The teacher’s evaluation is approached
to be broad and inexact. The most stigmatising feature of the holistic assessment is
unreliability. However, for teachers, it is the easiest method to be applied in written language
assessment.

To carry out such a survey a set of data and diverse methods have been used. The data
exploited in this work come from a corpus of linguistics examination papers of the second and
third year English LMD students at the University of Sidi-bel-Abbes during the academic year
2013-2014, and from some linguistics teachers’ interviews about the assessing writing
practice. The methodology used in achieving the present task is a qualitative comparative
analysis based on a psychometric /atomistic approach.

The present study includes four chapters. One of them is dedicated to academic researches
and theories about assessing the writing aptitudes, and three other chapters devoted to the
description of the study, the practical and comparative analysis. The first chapter is meant to
shed light on the study as a whole. It is devoted to the research aims, problematic, hypotheses,
the main and secondary populations of the study, the research tools namely the written tests
and the teachers’ interview, and the methodology used to accomplish the task. The second
chapter is an attempt to highlight theories and literature behind the issue under investigation
namely the writing assessment. It encompasses the objective and subjective testing of the
written production, and the essay assessment approaches namely the holistic approach and the analytical assessing method. The third chapter is devoted to scrutinizing the findings of the research tools mainly the two written tests and the teachers’ interview. It displays the researcher’s assessing procedure, the students assessing procedure, the sampling frame, the impressionistic assessing procedure, the essays’ samples analysis and the teachers’ contributions (Multiple scorings) along with their answers to the suggested interview. The last chapter is the researcher’s attempt to underline the effectiveness of the analytic approach to the writing assessment. Such an approach could solve the issue of the essay’s assessment reliability and objectivity since it may bring changes in the attitudes of EFL teachers as regards to the provision of the writing assessment so as to adjust, refine and enhance their assessing methods, techniques, and practices. The chapter also points out the possibility of introducing the computer in the writing assessment procedure. Today a great deal of automatic writing assessments is available. Assessing through the Computer technology is today a reality. It can be more effective and more reliable than any human assessing process.

All in all, the researcher’s assumption is that EFL learners encounter many difficulties when trying to develop an essay. Those difficulties are due to the ineffectiveness of the writing teacher’s feedback. And such ineffectiveness is due to the unreliability of the assessing procedure as a whole.
Chapter One

The Detection of the Complication
Chapter One The Detection of the Complication

1.1. Introduction

1.2. The Implementation of the LMD System
   1.2.1. LMD Features and Objectives
   1.2.2. English Language Teaching Within the LMD

1.3. Competency Based Approach
   1.3.1. Characteristics of CBA
   1.3.2. The Implementation of CBA in Algeria

1.4. Teaching Writing through CBA
   1.4.1. The Writing Deficits
   1.4.2. Writing Feedback
   1.4.3. Responding to Learners’ Writing

1.5. Assessing Writing
   1.5.1. Analytic Assessing Method
   1.5.2. Holistic Assessing Method

1.6. The Research Type
   1.6.1. The Research Purposes
   1.6.2. The Research Problematic
   1.6.3. The Research Questions
   1.6.4. The Research Hypotheses

1.7. The Research Population
   1.7.1. Main Population / Third Year Students
   1.7.2. The Main Population Linguistics Syllabus
   1.7.3. Secondary Population /Second Year Students
   1.7.4. The Secondary Population Linguistics Syllabus

1.8. Data Collection Devices
1.9. Secondary Population’s Written Test

1.9.1. The Students' Contributions to the Assessment Procedure

1.9.2. The Selection of the Written-Papers

1.9.3. The Selection of Students as Assessors

1.9.4. The Assessing Procedure:

   1.9.4.1. The First Procedure: Self-Assessment

   1.9.4.2. The Second Procedure: Assessment Exchange

   1.9.4.3. The Third Procedure: Multiple-Assessment

1.10. Primary Population Written Test

1.10.1. The First Tentative Assessment: Holistic Assessment

   1.10.1.1. The Researcher’s Assessments

   1.10.1.2. The Teachers’ Contributions: Multiple Assessments

1.10.2. The Second Tentative Assessment: Analytic Assessment

   1.10.2.1. The Researcher’s Assessment: The Threefold Assessment

   1.10.2.2. The Teachers’ Assessments

1.11. The Teachers’ Interview

1.12. Methodology

1.13. Conclusion
1.1. Introduction

Assessing EFL learners’ answers to an essay question has become one of the main nuisances that is still a subject of enquiry. The assessing procedure is subjective and entirely holistic in the sense that it seems impossible to control the marks that would be given. The teacher’s assessment is rather affected by personal opinion and prejudice which may result in rates unreliability. Such a factor is still a problematic in testing and assessing written language. On the other side, objective test such as cloze test items are rather systematic and analytic in a more accurate sense. The assessment is always reliable and trouble-free. The teacher’s assessment is, at some extent, based on a countable process. No judgement is needed on the part of the classroom teacher or other scorer who may be invited to have a word about the learner’s answer. The marks allotted remain static from exam paper to another, and from an assessment setting to another. This is due to the fact that the answer needed on the part of the students is either right or wrong. No space is left for the learner’s creativity and interpretation. However, such kind of assessment will reflect neither the learners’ lacunae nor their level of proficiency in writing.

The distinction between testing devices of a written language contributed greatly to rise out the unreliability of the holistic procedure which is encountered in assessing an essay. The identification of the problem that the foregoing study tries to resolve needed considerable data from both second and third year students. The results of the different exam sessions that had been set forth to both subjects, first year and third year students, were considered as indispensable tools to determining the real existence of the serious problem that hinder the assessing procedure. Such a fact had been proved by the contributions of the teachers who were begged to express their views about the existing phenomenon.

The identification of the problem was a key answer that impelled the classroom teacher to adopt another strategy. The first attempt in assessing put forward clear fingerprints on the
testing of essays which needed a practical and useful remedy. The inclusion of the Analytic method as a second attempt had been adopted in order to concretises and objectify the assessing procedure. The answer to such an attempt could be proved or disproved by the contributions of different hands, and minds whose interpretations of the students’ responses might not be affected by impressionist judgements.

However and before deciphering the results, it advisable to highlight many points concerning the educational setting. Thus, this chapter starts by describing the new reforms implemented in the Algerian university characterised by the introduction of the LMD system and the adoption of Competency-based Approach (CBA) as a key approach to teach mainly writing. The present chapter also introduces, in its second part, the research type, purposes, problematic, hypotheses, population and the different tools that were used in order to identify the problem and to find out a solution that would be more formal and practical in testing and assessing the EFL learners’ written language.

1.2. The implementation of the LMD system

Historically speaking, the Algerian University has experienced a great deal of reforms since its creation immediately after the independence. The attempts to adapt the university to the national and international new requirements have been marked by different reforms, including that of 1971, which profoundly restructured the scheme of the higher education and scientific research in Algeria. Such a reform was meant in particular to meet two challenges: first, the creation of technological and scientific elite that may realize a better development for the nation, and second the Arabization of higher education.

As matter of fact, the Algerian university knew a huge quantitative progress mainly during the last twenty-five years. Significant infrastructure and human expansions have flourished within the higher education field. According to the Algerian Ministry of Higher Education and Scientific Research, the sector encompasses 1.930.000 enrolled students, ninety-one
institutions, 44,000 teachers, 27,000 researchers, and 33,000 as doctorate students. However, such an unplanned immense evolution was not without a tax to pay. This significant and rapid growth has generated many constraints. A sequence of troubles has appeared and led to an astonishing decline of the learning and teaching quality at the university. Furthermore, a severe inconsistency between community needs, and university production could easily be noticed.

In order to set up these problems, several committees have been established. Thus, many conclusions were pointed out to determine the shortcomings of the whole university system and to suggest some renovations. These committees mentioned the deficiency of the university staff, the ineffectiveness and inappropriateness of both university syllabuses and themes, and more importantly the giant gap between the university and the socio-economic requirements. Accordingly, decision makers blamed the old organism put into service since the independence. The classical structure, based upon four years to get the licence, two years for magister and four years for the doctorate, seemed no longer to comply with the recent challenges imposed by the new international political, economical and social trends.

Consequently, there was an urgent need to reshape everything in the Algerian University to meet the new global patterns and trends in such a field. It is in this perspective that the new reform fits higher education. The implementation of the European instructive scheme branded as LMD (licence, master, and doctorate) system was more than just a choice in 2004. The implementation of the LMD system in the Algerian university was meant to bring an adequate and effective assessment and evaluation. Moreover, the new reform aimed at enhancing the reliability and the effectiveness of the whole process of learning and teaching.

However, the main concern, in this humble chapter, is about the new dimensions brought by the LMD system to the learning and teaching of English as a foreign language. In more
price terms, the chapter may shed light on the impact of the LMD system on both teaching and assessing the written language of the EFL learners.

1.2.1. LMD features and objectives

The Algerian Educational system has witnessed a kind of a premeditated but a very slow progression as regards to its teaching methodologies and curriculum content development. Series of renovations have been introduced to adapt the higher education to the on-going globalization process.

The implementation of the LMD system in the Algerian university is viewed as a shift towards Globalisation since the so-called model has been approved by the world most powerful and developed countries either in Europe or in America. Decision makers were looking for how to renovate the Algerian universities through well structured and founded form of a long-term education. LMD was hoped to allow a real re-foundation of the whole university programmes and approaches which have become useless and out of date to achieve progress and to make such successful Algerian engineer, teacher, doctor, lawyer or others.

Because of its freshness, it is worth to shed a slight light on the LMD, its main components, features and objectives. It includes three key different stages: the Licence after six semesters of studies, the Master degree after four semesters and finally the Doctorate after six other semesters. Before putting into practice such kind of higher education, a lot of debates and discussions have taken place at many regional and national seminars all over the country. The implementation of the LMD system aimed at introducing an overall innovation to the whole Algerian university scheme to cope with the modern universities all over the world.

The new thing that the LMD may have brought is the type of evaluation and the criteria to allow the student pass successfully his academic year. The new sort of assessment is based on the system of “Credit accumulation” i.e. every student is required to get certain amount of
credits. However, in case the student got the necessary total of credit but he missed some subjects, he may get to the following semester and have an extra opportunity to pass the previous subjects. The LMD system was put into practice during the academic year 2004/2005. It was such great responsibility that many universities didn’t apply immediately the LMD till they got some ideas about from the other pioneer universities like those of Tlemcén in the west, Algiers in the centre and Constantine in the east.

1.2.2. English language teaching within the LMD

English has become the language globally used for human contact. Currently, no one may escape the general rule of being involved in any scientific, medical, economical, political, cultural, social and even business matters where English holds a central significance. Algeria seems to be aware of those essential prerequisites for English. The language map of Algeria has witnessed a momentous through such huge shift from French to English as a subject in the educational curricula, or the escalating amount of the EFL learners in the Algerian universities. However, the history of foreign language teaching in general and Teaching English as Foreign Language in particular has been and will continue to be the history of controversies, debates and researches.

In the newly introduced system, the EFL learners must go through six semesters to get the license. During their third year, they are invited to choose either to write a thesis about a theme suggested by the supervisor or to embark on a teaching training concluded by a written report. Shortly after having the Licence Degree, students may have access to the Master studies for two years, and then to the Doctorate studies for three more years. On the other hand, the LMD teachers are restricted to certain pedagogical guidelines. They are responsible about the content to their students’ needs. The first year LMD English syllabus focuses mostly on language skills. During the second and third years new subjects are introduced according
to the Licence title. Each Licence functions upon a specific sphere of study most of the time related to the particularities of the area the university is located in.

LMD system in Algeria is seen as a step towards a new teaching outlook due to the great necessity to rectify the Algerian educational system at universities and coping with the new requirements of global trends of higher education. This reform is mainly set to give further opportunities to both teachers to enhance their teaching and students to broaden their learning. On the one hand, teachers are assisted to present training courses fully based on an active path, on the other hand, it offers the students the possibility to go through a real effective training. In order to make the LMD architecture more beneficial, a number of measures have to be adopted at the level of curriculum design and integration. The Algerian Ministry of Higher Education had to rely on innovative teaching methods to recover the linguistic competence and to reach the following goals put forward:

- Guarantee a high teaching quality.
- Being opened to the universal development, technology and sciences.
- Encourage global and international diversity.
- Set a balance between the University and the world through all possible interactions.

The Algerian educational system adopted the Competency-based Approach (CBA) in an attempt to achieve the planned aims.

1.3. Competency Based Approach

Historically speaking, CBA is considered as an old idea that was first introduced in the late 1960’s in the United States of America and evolved through the implementation of other professional, educational and vocational training programs in Germany, UK, and Australia too in 1990’s. Some commentators argued that the theoretical conceptions of the CBA lie totally on the human psychology and more precisely in the behaviourists’ models that are about making relationships between competencies on the basis of the performance. Frederick
Taylor (1947) is the most important historical figure since he is one of the founding fathers who first trialled the CBA as it applied nowadays.

1.3.1. Characteristics of CBA

The main characteristic of competency-based approach is its focus on the design and evaluation of competencies or outcomes. CBA concentrate more on the ability to apply knowledge to perform practical tasks and to fulfil workplace roles rather than emphasizing theoretical or book knowledge. A competency is the ability of the learner to put skills and knowledge into action. The competency-based approach stresses how to obtain life coping skills without losing insight on improving the language to act upon these skills. It is founded upon hypotheses of learning which state that for effective learning to take place, students have to recognize that what they are learning would develop their lives. The approach has been designed and put into practice in the USA to assist immigrants and expatriates learn English and life skills at the same time.

By 2005, EFL teaching methodology witnessed the implementation of the CBA in an attempt to make the educational product (i.e. graduated students) complies with the world of work, thus creating a kind of harmony to strengthen the society as a whole. Students are supposed to be provided with both the skill and the knowledge to create a competency to afford a better life. Such a competency could be defined as a savoir-faire or a know-what combined with a know-how. In more exact terms, it is a formula to integrate and put into practice a set of abilities and a certain amounts of knowledge that will be used successfully and effectively in a variety of problem-solving conditions or circumstances that have never taken place previously, i.e. a competency continues throughout and beyond the school curriculum.

Competencies in general are gained after experiencing different tasks in order to achieve excellence in practicing and performing certain skills, making people up dated to their world
There has been however a growing interest on the educational systems about implementing new approach based fully on competencies. Relating this to our context, the CBA was adopted in teaching English as a foreign language in order to prepare students to be skillful, capable and able to be competent in their real life issues. The competency-based approach (CBA) to teaching English is similar to the communicative teaching approach, that almost all the Algerian teachers are often familiar with. It is however one of the commonly known approaches which emphasizes generally on the useful skills, knowledge and abilities focusing not only on the input but rather on the results or the outcomes. It addresses what the learners are expected to learn about. It refers to an educational movement that advocates defining educational goals in terms of precise measurable descriptions of knowledge, skills, and behaviours students should possess at the end of a course of study. 

Richards and Rodgers (2001:88). Besides, the CBA as a functional approach that evaluates learners’ mastery of skills, abilities according to their performance, as it has been defined by the US Office of Education as “…a performance-based process leading to demonstrated mastery of basic life skills necessary for the individual to function proficiency in society.” (US Office of Education 1971).

1.3.2. The implementation of CBA in the Algeria

Generally speaking, the CBA is seen as a cognitive approach that is based on Bloom’s taxonomy of Education objectives namely the cognitive objectives. CBA encompasses everything that deals with affect, emotions, values, and body gestures. CBA is a learner-centred approach in which the student is required to go through a gradual process where he has to question his personal connections, revise his acquired knowledge and compare his gained information with the mates in the classroom. Just as an attempt to make a balance between his prior knowledge and his newly required one. On the other side, the teacher plays an effective role in facilitating the task of learning where he would be less directive (giving
information and answers) and instructive when designing tasks to meet his learners’ needs, and in return learners take part in an active manner in order to express themselves and comprehend the others’ messages. The teacher acts as a researcher, a careful observer, and a listener to his learners that obviously may help him learn more and more about the learners’ preferences and styles to learn.

Modern societies are privileged by their progression and rapid changes not only in the field of education but rather in all spheres of real life. In this vein, firstly learners are supposed to be effective integrators’ with the environment and being able to communicate easily with people they meet from different cultural and social groups. As a result they will be responsible for rolling and managing their own issues. Due to the shortcomings of the communicative approach, the CBA is hoped to help the Algerian learner be aware of the opportunities given to him in finding solutions to his problems and improve his ability to act according to the vision of the world that he is growing in.

Competency-based approach to teaching English language aims at motivating the learner. CBA aids the learner recognize the value of the language and the skills he is learning, express himself freely when using English and exchanging ideas, and take an active role when being involved in any learning process. It is very tiring for teachers to be the centre of attention the whole session. So, through the competency-based teaching classroom, learners take a more active role in stead and the teacher is not required to be the centre of attention all the time. Moreover, it is more important for the teacher to own the whole control, freedom, and ability to be creative in a competency-based classroom since he is able to use approaches, methods and techniques of teaching that better fit himself and his students, and achieve the immediate objectives of the session.

In short, Algeria’s step in implementing such a reform comes just to enable people and mainly EFL learners to reach a high level in terms of requiring new knowledge and
competencies. This approach therefore, is a kind of a revision of the teacher’s attitudes, teaching profession and knowledge.

1.4. Teaching writing through CBA

According to Walters (1983:17), writing is “the last and perhaps most difficult skill students learn...if they ever do. Recently, teaching writing focused on the linguistic forms. Speech was the main, skill and writing was viewed as a way to support the oral model of the language. Raimes (1983:6) sees writing "served to reinforce speech in that it stressed the mastery of grammatical and syntactic forms." Writing was usually a question of completing, reproducing, replacing and converting, in brief writing use to deal with the structure of sentence. Writing was meant to check and test the accurate use of grammar conventions. In other words, writing considered the forms and structures composed by the EFL learners. Learners were allocated excellent marks, whenever they succeeded in producing error-free written texts. Writing was seen as a substandard skill to speaking i.e. writing should strengthen the speaking skills. Widowson (1978:116) points out that “they need to pay more attention whatever to what the sentences mean or the manner in which they relate to each other.” EFL learners used to be assessed according to how correct their grammar was. Writing drills were no more than converting manuscript from past to future simple tense for instance, in altering the entire masculine to feminine, and so on.

However, educationalists and scholars gradually recognized that the emphasis on the product neglect the process of writing. Furthermore, the students were not authorized independence to produce their personal writing. Researchers perceived that the attention was paid to structure of writing not to the way how learners may develop a written paper. Actually, the teaching of the writing skills is not a method to underpin speaking language knowledge. Learning to write is not a matter of learning how to convert and transform others’ writings, it rather goes beyond towards generating words, sentences and expressions. Writing
is a question of searching, finding and transmitting thoughts to others, as Shaughnessy (1977:234) perceives it: "the record of an idea developing. It is a process whereby an initial idea gets extended and refined."

Currently, the EFL learners, in any writing class, do not go through identical sorts of drills. The EFL learner is viewed, today, as a creator of any written piece. Therefore, the process approach with innovative classroom activities and assignments has emerged. The use of diaries, brainstorming, drafting, cooperating, revising, editing and concentrating on both content and form are the new features of the writing process.

The Algerian educational system makes no exception to the rule. New reforms were introduced to the teaching of writing. Such reforms are marked by a shift from the product approach towards the process approach. Moreover, the final objective of teaching writing has become that, at the end of any writing sequence, the EFL learners should be able to produce a piece of writing of about twenty lines through the competency-based approach. The writing skills are indispensable for the success of any EFL learner for the simple reason that in almost every subject the learner is supposed to express himself in a written form during that famous examination at the end of each semester.

Writing, as Bloomfield (1990:101) fixes, is not only a method to record language through the use perceptible symbols but also a communicative exercise and a process to generate a meaning. Moreover, Wilson (1994:16) sees writing as "a creative process of immense perceptual linguistic and cognitive complexity in which meanings are made through the active and continued involvement of the write and the unfolding text". Writing is a too demanding complicated drill. Writing usually requires the learner to investigate many issues, to improve his faculty to assume, consider and feel. Byrne (1979:14) thinks that writing means transforming our thoughts into language. It is a very complex skill that requires both physical and mental activity on the part of the writer. Writing is a determined choice and association
of information. Writing is a tool to deliver in a beautiful way our opinions, emotions, and experiences. Judy (1986:39) and Laver (1980:56) say: "Students' writing should be motivated by their feelings about and responses to a topic with which they have some experience".

Besides, it is a recursive process which involves: brainstorming, drafting, editing and writing. In fact, a good writer thinks, writes, rewrite, and writes again. Advising writers, Nicolas Boileau says: “Bring your work to the workshop twenty times. Polish it continuously, and polish it again.” He adds: “Four every four words I write, I strike out three.” Boileau points out the significance of the writing process to achieve a better performance. Writing is now seen as a communicative process. It represents as Widdowson (1983:4) puts it: “An interactive process of negotiations of meaning." It is not a private matter. The feedback of other learners to a given written piece may assist the writer find out thoughts, edit and enhance the organisation of sentences within the paragraph and the paragraphs within the text. Indeed collaborative writing could solve many writing–related problems.

In this respect, CBA seems of a paramount significance in enhancing writing effectiveness. Alison Oswald (2012:54) assumes that "competency-based learning is motivated by the idea that the classroom should prepare learner for real life. Learners acquire the language, but they also develop skills and strategies that will help them complete the kind of tasks they would do in real life." Everything the learner writes is the product of a step-by-step process. The process means prewriting activities, and the students are supposed to write drafts, make revisions, then come up with a final product. The process oriented approach includes multiple stages such as: pre-writing, writing, and rewriting. Once the rough draft is created, it is polished into other drafts with the help of the other readers (teacher and peers), then final editing is followed. Writing activities help the students to be familiar with patterns of language they need for effective communication i.e. features such as linking devices, forms of
language, and vocabulary. On the other hand, the teacher develops an approach to writing for communication. Gradual explanation and exercises’ enable the students to accomplish, with a sense of achievement, and through the process of writing a final product. As Perl (1980:368) points out, “When we are successful at this process, we end up with the product that teaches us, that clarifies what we know (or what we know at one point only implicitly) and that lifts out or explicates or enlarges our experience.”

Thus, writing is both process and product, and the teaching of writing does not involve only the teaching of its mechanics (grammar, vocabulary, spelling...) but also of its communication and use. Learners need plenty of opportunities to write, draft, rewrite, and redraft. In the way towards the final version, learners have to move around thoughts, arrange them, build, and rebuild patterns to create an essay which includes accuracy of form, suitability of style and coherence. Therefore, teachers who try the process approach with their students help them develop good strategies for writing, and create a stimulating atmosphere in the classroom. Grammar and vocabulary are also important components of a good written piece. Savignon (1993:43) sees grammar as an important component since “learners seem to focus best on grammar when it relates to their communicative needs and experiences.”

The writing learners need to examine samples of writing, practise the forms of language, get help during the writing process, analyse the structure of given texts, practise pieces of communication in order to achieve a better writing performance. It is important for them to see that controlled exercises have their place in writing.

Consequently, an eclectic approach to the teaching of writing will help EFL learners in the process of writing. In fact, grammar, syntax, mechanics of writing, organisation, word choice, content, the process, audience, and the purpose are all components in the composing process, for they help to communicate ideas.
1.4.1. The writing deficits

The changing tendencies in writing theory have been accompanied by changing views regarding deficits. Raimes (1991) recapitulates the new trends saying that with controlled compositions methods, errors were avoided. With the shift to the process approach to writing, risk taking was encouraged and errors were tolerated, and in some cases, ignored. Raimes (1991) recognizes that English writing teachers cannot close the eyes over the grammar inaccuracy but envisage that it will vanish over time. She underlines that many advanced EFL learners have to recover the correctness of their written English within merely a semester or two. They are challenging, to a certain extent, natives, so they should enhance their aptitudes for treating their deficits in writing.

Accordingly, Raimes (1991) and many other EFL researchers propose that instructors boost and valorise the establishment of self-editing talents in their learners. The responsibility to correct errors has shifted from the teacher to the students. Teachers’ correction seems to have insignificant influence over recovering students’ writing, particularly considering the time teachers ought to allot to the correction sessions if there is any. Moreover, it has been argued that teachers’ correction may create certain dependency, even though some researchers consider a specific amount of reliance to be practicable for intermediate or untrained writers. Learners actually allocate a great deal of attention to grammatical hindrances whenever they have the opportunity to localise themselves in someone else’s writing or in their own, instead of relying on the teacher to categorize them. Likewise, they deal with mistakes more efficiently in situations where only one draft on a writing assignment is required. They also pay more attention to teacher’s feedback when they are in composition classes that employ a multiple-draft approach rather than the single draft approach. Raimes (ibid) and other writing specialists suggest that errors could be approached analytically, on the basis of occurrence and language samples as well as their
intensity and gravity. They assert that teachers have to be able to spot patterns in students’
errors and teach with fitting strategies. The main concern (Zamel, 1985) is that teachers
adopt a process rather than using a random approach and that the process effectively
addresses the meaning loss incurred with the error.

1.4.2. Writing Feedback

The current inclination towards a more learner-centred approach in FL learning/teaching
has led to a more challenging role for both teachers and learners. The most important change
is that the teacher is no longer “the distributor of knowledge” or the dispenser of sanctions
and judgements. Furthermore, this change has provided learners with more responsibility and
contribution to the learning process. Thus, the teacher’s role is less prescriptive and less
overprotective because learning is seen to be a collectively shared practice. This combined
effort assumes a more involvement of the learner in the learning process (Nunan 1988).

Today’s learner differs deeply from the classical learner for he is a lively contributor who
accomplishes his task and builds up a consciousness of himself as a learner. Within this shift,
there has been a key change in the hypothesis and investigation over writing. More stress is,
now, put on the process of writing. With the conventional conviction of error-free writing the
teacher had to put off the happening of errors at all price. However, feedback that focused on
errors does nothing to assist learners in creating and surveying ideas. This feedback provides
no reflection upon reader-based discourse, let alone other basic features of the writing
process.

With the process approach, both teacher and learners are collaboratively involved in
producing a piece of writing. They know that writing is recursive and a cyclical process
during which learners go back and forth on a whole gamut of expressions and structures. This
conviction argues for an entirety dissimilar feedback system; it gives more importance to
how discourse is produced through finding and negotiation on both content and form.
The traditional product-oriented feedback focused mostly on form. It highlighted grammatical rightness and abandoned other aspects of writing process like the generation of ideas, the construction of structures and the building of meaning for all anxiety was about language exactness. Zamel (1985) demonstrates that the traditional feedback procedures fail to make correct grammar automatic as well as to adequately stress particular aspects of writing. Moreover, Krashen states that:

If the student-writer is able to consciously learn all the rules of punctuation, spelling, grammar, and style that linguists have discovered and described his reward should be a Ph.D. in linguistics. Unfortunately, this will not guarantee him writing competence, since so much of what good writers do routinely and subconsciously remains to be discovered.

(Krashen, 1984:25)

This sort of feedback is more similar to a marking session, which may make learners feel bothered or even hating the writing class, because no chance is provided to redraft the written production. Many evidences confirm that this feedback does not have an important consequence on boosting learners’ writing aptitude. So superficial and insignificant correction does not considerably relate to writing quality, and the correction of errors by the teacher could simply assist the learners’ affinity to focus only on sentence-level troubles, losing insights upon the discourse as a whole. In other words, when learners are preoccupied by grammatical accuracy they can scarcely concentrate on generating ideas and conveying them suitably in a written discourse.

Undeniably, through writing the learner is engaged in a complex procedure that necessitates both the content and the procedural knowledge that facilitate the management of this content and the construction of a piece of writing of a specific kind. Educationalists agree that in any skill acquisition, feedback is of a supreme significance. It is regarded as a way to provide data and support for revision. This feedback encompasses comments,
remarks, and notes. Briefly it is a method to supply information about the writing performance that is used as an input in the learning process. Consequently, feedback could achieve the following functions:

- Feedback is persistent, i.e. feedback and reflections have to be included within the writing process as recursive and cycling events.
- Feedback is an input for editing, i.e. both feedback and responses are used as input for revising and proofreading.
- Feedback is not rating, i.e. providing feedback is not giving marks. The feedback and responses have not to be delayed to the last stage of the writing process.
- Feedback is supplied as appropriate, i.e. the learner could receive different feedback and responses at different stages in the writing process as appropriate.

Writing teachers could try a whole gamut of techniques to make the feedback more effective. The teacher and the learner, for instance, may meet on one-to-one basis as editor and writer to discuss the draft text. This learner self-monitoring technique is the best way to provide feedback for it encourages learners to look critically at their writing on the one hand, and the teacher may provide tailor-made feedback to individual learners on the other hand. Yet, it remains difficult to achieve because of time and classroom arrangement. Most teachers confess that conferring with learners is the finest and the most efficient way to provide them with individual writing attention. Both teachers and learners can develop the kind of relationship that promotes writing proficiency. These conferences have to be as frequent as possible depending on the time allotted to writing instruction. Teachers can devote twenty minutes to meet with learners focusing on one or two particular writing problems. Studies have shown that talking with learners is the best way to help them improve their writing ability.
1.4.3. Responding to Learners’ Writing

Thanks to the availability of a wide range of writing approaches, teachers often have a number of ways to react to learners’ writing. The response to writing is potentially one of the most influential tasks in writing class. Teachers can spot on errors, specify the amount and the nature of errors, and comment on content. Responding to learners’ writing is, then, a thorny matter; it requires the teacher to allot the paper a score, write a remark and correct mistakes. However, the act of responding often marks the end of the interaction and learners is forced to stop there. The teacher’s response illustrates and point out details and guidelines on what to do next, and extraordinary sessions are devoted to cure. Learners are invited to rewrite after collective and individual correction. The teacher’s response may be of huge assistance to enhance learners’ writing proficiency.

When learners submit a piece of writing many aspects are to be assessed: if the ideas or proceedings are important, well arranged, simple to follow and enjoyable to read; and whether grammar, vocabulary, spelling and punctuation are of an acceptable standard of accuracy. Among the various writing components, teachers usually find themselves relating to forms in their feedback, because mistakes are easily diagnosed and corrected than content and organization; they are difficult to ignore. Furthermore, learners want their mistakes to be corrected. (Penny Ur, 1997)

Assessing procedures can make response to any piece of writing a prolongation of a writing course; it is an indispensable part of the process of teaching writing. Teachers can make of learners’ errors a learning situation remedying the weaknesses and improving their writing proficiency. However, most language teachers consider the summative assessment or the exam as the final step in the process of teaching. Such a common slip is really impeding both learners and teachers from an effective opportunity to redress and reset the whole
pattern of learning/teaching process. Then, the sequence of classroom writing follows somehow this pattern:

![Diagram of Writing Class Pattern (Adapted from Raimes, 1983)](image)

**Figure 1.1. Writing Class Pattern (Adapted from Raimes, 1983)**

Writing assessment is practical chiefly as a means of enhancing learning. Writing feedback serves to revise existing writing curricula and/or plan programme for individual as well as drive pedagogy and improve teaching-learning environment. The writing assessment and feedback demonstrates methodical validity that encourages classroom practices. Responding to learners’ writing has always been an important way to efficiently individualize learners’ writing teaching. Most teachers report that they spend a great deal of time responding to their students. First drafts naturally last longer to respond than final drafts. A successful feedback to any writing depends on a diagnosis of the paper’s hindrances like the thesis, the structure, the paragraph and the style. After diagnosing students’ writing hindrances, teachers try to design four different brands of responses to their learners’ written pieces: facilitative, directive, corrective and evaluative. Including all the previously
mentioned features, the teachers’ responses and feedback could achieve better effective impacts on the learners’ writing achievements. Facilitative reflections are crafted observations as questions that motivate learners to consider their ideas and their expressions. The idea behind the facilitative response is that students learn best to write and develop revision strategies as well as they retain responsibility for their own writing process.

On the other hand, teachers sometimes think that facilitative responses are not suitable for particular learners’ assignments. They agree that some learners need more explicit writing advice, i.e. they want to know where they went wrong and why. Therefore, directive responses are more fitting and effective. Teachers may, for example, ask learners to remove paragraph, omit sentences or change a word. Such directive responses, like “omit” and “remove”, are instructive when they are escorted by explanations, for example, “omit” because of irrelevance, redundancy, sentence structure, paragraphing and so on.

The third kind or the corrective response is typically copy-editing notes, which points out different writing errors like in grammar and vocabulary by labelling errors using specific writing symbols, circling and marking them according to a particular code, correcting or re-writing (with the aim of modelling a correct eloquent style). Finally, they ask learners to correct their errors in particular writing courses.

The last sort of response to consider is the evaluative response, a response that usually includes a rate. This rate reflects, to some extent, the writing content. Some teachers prefer giving two grades: one for the content of a paper, another for its writing or style. Some teachers do not grade the first drafts at all arguing that grades distract learners from the real process of exploring an idea. Whereas others grade the first drafts of papers more harshly than they do with the final drafts, because they hope that learners will improve as they revise.

Different writing responses work best in different writing situations. But teachers believe that facilitative approaches to responding to learners writing bring them to think critically
about their own work and places writing responsibility and thinking squarely on the learners' shoulders. On the other hand, teachers think that their responses can be either remarks put in the margin of a paper, or a summary of remarks made at the end. It is worth noting that teachers have to make different kinds of comments in different places of the paper: corrective, facilitative questions and praise for an idea are likely to be found in the margin. Teachers often begin their closing comments with praise of something well-done like an argument or readable of prose style. Then, they ask learners to consider certain points more thoroughly. Next, they comment on structure, organization and logic. Finally, they address matters of grammar and language usage.

1.5. Assessing Writing

Having responded to the learners' assignments is only half of the task. Assessing them is equally challenging; assessing has always been very tiring and time consuming. Assessing and scoring writings are the most backbreaking; it involves more hard work and attentiveness for more reliability and objectivity. Assessing writing ability of E.S.L and/or E.F.L learners has become increasingly of great importance in recent years, because such a response is used for a variety of administrative, instructional and research purposes. The issue of deciding the use of a single score for a given written text or different features of the written text is scored separately. Among others, two types of assessment procedures have mainly been discussed: the analytic method the holistic method.

1.5.1. Analytic Assessing Method

For those who view writing as a set of separate skills rather than as an overall generalized ability, they prefer the analytic method. This method involves the separation of various features of a piece of writing into many different components. It is based on the idea that individual writers may have strengths in some areas and weaknesses in others. Traits of good writing are broken down into categories like organization, development, awareness of the
audience, mechanics and coherence. Thus, analytic method gives more detailed information about writing performance.

Many writing specialists prefer the analytic method over holistic for a number of reasons. They say that the analytic offers more practical and indicative information about the learners’ writing aptitude. In other words, it demonstrates and spot their strengths and weaknesses, and thus permits teachers and syllabus designers to adapt directives proper and more closely to the learners’ needs. It is more fitting for E.F.L learners who are more likely to show an obvious profile across different facets of writing. Some learners, for instance, may have outstanding writing talents in terms of content and organization, but may hold inferior mastery over grammatical issues.

Nonetheless, the analytic method spare much more time than the holistic one, for teachers are required to make more than one decision for every writing sample. Moreover, measuring the quality of text by tallying accumulated sub-skill scores lessens the interconnectedness of written discourse, and gives the false impression that writing can be understood and fairly assessed by analyzing autonomous text features (Hillcocks, 1995; White, 1994). Consequently, these broken down smaller units affect the assessment as a unified whole. Finally, qualitative judgements about coherence and style are not easy to be assessed even by experienced teachers (Hamp-Lyons, 1991).

For Huot (1990), however, the analytic method tends to have higher reliability than holistic. It is more useful for assessing various dimensions of learners’ abilities, but it is also potentially more valuable for prescribing educational interventions for individuals. Further, this kind of assessment can lead to reform the curriculum if most learners exhibit similar patterns of deficits. All in all, it is worth noting that no assessment procedure is suitable for all purposes even with guidelines and set criteria; both methods can yield unreliable
information. Therefore, the choice and decisions need to be made within the context, the situation and the purpose.

It is sometimes advisable to disconnect each characteristic and sub-skill of writing for scoring reasons. This method is more objective and reliable. The scorer may set up a grid of different writing components like grammar, mechanics, fluency and relevance, ease of communication and the content related to the topic. The analytic method is valuable for the teacher to determine where learners need to improve and how they have obtained their marks. It attempts to evaluate components of a piece of writing. The scorer may use the point-off method, i.e. the learner loses points for errors that occur in his piece of writing. In order to avoid failure for repeated errors other systems like a unit-off for one or two errors, or give different errors weights, like the one suggested by Madsen (1983:121).

![Graph 1.1. Madsen’s Analytic Scoring Method](image)

The analytic scoring is so useful for diagnosis purposes and a more objective evaluation. It is preferable to react properly for teaching-learning commands. The teacher may consider the following writing items and attributes as displayed in the table to grade the written language, as suggested by the Oregon Department of Education, in the USA.
Table 1.1. Scoring Model Adapted from Oregon Department of Education 2005

<table>
<thead>
<tr>
<th>Conventions</th>
<th>Vocabulary</th>
<th>Sentence structure</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Above Standard -A-</strong></td>
<td>-Accurate and right use of vocabulary</td>
<td>-Use of sophisticated sentence patterns: simple, compound and complex sentences</td>
<td>-Creative original ideas -Good supporting ideas -Precise supporting details</td>
</tr>
<tr>
<td>Ideal use of punctuation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Capitalization</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use margins</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intended paragraphs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>At Standard -B-</strong></td>
<td>-Acceptable use of vocabulary</td>
<td>-Simple and complex sentences are used</td>
<td>-Thoughtful and clear ideas -Relevant supporting ideas</td>
</tr>
<tr>
<td>Adequate use of punctuation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Capitalization</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use margins</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intended paragraphs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Approaching Standard -C-</strong></td>
<td>-Poor use of vocabulary</td>
<td>-Simple sentences are used</td>
<td>-Inappropriate ideas -No evidence -Very vague -Irrelevant details -Confusing</td>
</tr>
<tr>
<td>Inconsistent use of</td>
<td></td>
<td>-Sentence structure is usually correct</td>
<td></td>
</tr>
<tr>
<td>punctuation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Capitalization</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use margins</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intended paragraphs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Below Standard -D-</strong></td>
<td>-Incorrect and wrong use of vocabulary</td>
<td>-Sentences do not make sense</td>
<td>-Impossible to assess the content</td>
</tr>
<tr>
<td>Incorrect use of punctuation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Capitalization</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use margins</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intended paragraphs</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The letters A, B, C and D represent grades allocated to written pieces in case the appropriate traits are found in the written paper. On the other hand and as far as writing performance appreciation is concerned, the following table displays an effective gamut of comments and impressions according to the availability of the corresponding traits, adapted from Moskal.
<table>
<thead>
<tr>
<th>Conventions</th>
<th>Vocabulary</th>
<th>Sentence structure</th>
<th>Content</th>
<th>Organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excellent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| -Accurate spelling
  -Effective use of punctuation
  -Correct capitalization | -Accurate words
  -Variety of words
  -Original expressions | -Effective sentence structure
  -Good command of sentence patterns | -Exact response to the topic
  -Details are precise
  -Ideas are carefully chosen | -Interesting introduction
  -Appropriate focus
  -Well arrangement of paragraphs
  -Related conclusion |
| Adequate     |            |                    |         |              |
| -Correct spelling
  -Proper use of capital letters
  -Correct punctuation that helps understanding paragraphs | -Words communicate the main ideas
  -A variety of words used to fit… | -Good control of sentence structure
  -Different types of sentences | -Thoughtful ideas
  -Relevant and specific supporting details | -Clear introduction
  -Appropriate transition
  -Adequate conclusion |
| Acceptable   |            |                    |         |              |
| -Limited spelling errors
  -Correct capitalization
  -A few punctuation errors | -Correct words
  -Some words do not fit
  A few repeated words | -Sentence structure generally controlled
  -Limited variety of sentences
  -Some laps | -Relevant details but general
  -Appropriate use of ideas | -Vague introduction
  Lack of some coherence
  -Artificial conclusion |
| Need improvement |            |                    |         |              |
| -Frequent spelling errors
  -Incorrect and inconsistent capitalization
  -Confusing punctuation | -Many repeated words
  -Limited and poor use of words (lack of variety) | -Lack of control of sentence structure
  -Limited use of language usage | -Irrelevant and vague supporting details
  -Some ideas are out of context | -Introduction lacks purpose
  -Lack of coherence
  -Poor transition
  -Unrelated conclusion |
| Fundamentally deficient |            |                    |         |              |
| -Too many spelling errors
  -Wrong use of capital letters
  -Absence of punctuation | -Words do not fit
  -Confusing words
  -Use of wrong words
  -Very limited choice of vocabulary | -Very poor language structure
  -Lack of prerequisites | -So limited that it is not possible to assess
  -No attempt
  -Wrong response | -Inappropriate introduction
  -No transition
  -Conclusion missing |

Table 1.2. Writing performance appreciations (adapted from Moskal 2000)

The teacher, nevertheless, has to justify any of the recommended appreciations by providing the suggested impressions and comments. Thus, the leaner may determine his deficits.
1.5.2. Holistic Assessing Method

Holistic method is meant to rate the overall ability level reflected in a given sample of learner’s writing. The writing is quickly read and judged according to a ranking scale that sketches particular standards. In other words, the holistic method is supposed to identify the whole core or value of writing products. This method involves no more than three (sometimes less than three) indicators awarding a single mark that is based on the total impression of the writing as a whole.

The holistic method, which yields one broad mathematical score, possesses the obvious advantage of speed making it more useful than other methods. More efficiency is another important consideration because it is more convenient since teachers are required to make one decision for each writing sample.

However, the key disadvantages of holistic method appear from the limitations of the single score, which gives useful ranking information but no details. The attributed score does not differentiate between the various writing components (syntax, vocabulary choice, organization, conventions and the like). The critical restraint of the holistic method is the lack of diagnostic information pertaining the learners’ strengths and weaknesses.

Carson & Nelson (1996) point out that despite careful labours dedicated to training scores there is always some degree of subjective prejudices involved in holistic method. These personal prejudices may be problematic when both the writer and the reader possess discrepant sets of cultural conventions and expectations.

The mark is extremely subjective opinionated judgement of the learners’ written performance. Judgements are often influenced by handwriting neatness as well as other factors that may significantly disfigure true assessment. Therefore, it is worthwhile for teachers to rest and resume work when they feel better. Yet, the holistic method is found more enjoyable, easier and faster than any other way of assessing. Hughues (1989:86)
confirms that “Holistic scoring (often referred to as impressionistic scoring) involves the assignment of a single score to a piece of writing on the basis of an overall impression of it”. The following adapted graph may illustrate better the difference between the analytic and holistic methods of writing assessment regarding the qualities and advantages of each.

![Contrastive Graph of Analytic versus Holistic Methods](image)

**Figure 1.2. Analytic versus holistic methods, contrastive graph**

*(Adapted from Weigle, 2002:121)*

It is so obvious from the graph above that analytic and holistic assessing methods are different regarding many aspects including their reliability, their practicality, their impact, their authenticity, and finally their validity.
This was briefly a description of the educational setting referring to the form and content of teaching, testing, assessing and grading writing. The researcher, now, would highlight every detail about the ongoing investigation.

1.6. The research type

The present work is an experimental study. It is a cross-sectional investigation. It was conducted in the department of English at the Faculty of Letters and Arts at the University of Djillali Liabes, Sidi-bel-Abbes, during the second semester of the academic year 2013-2014. It is based on cautious hypotheses. These hypotheses are the answers of a series of questions which deal with an actual dilemma that exists in scoring and assessing the EFL learners’ written production particularly the essay. The hypotheses are, therefore, tested and followed by observation of results. The researcher enlisted services of some of his colleagues who worked on the same ground. Some students had also taken part in determining the efficiency of such a type of research, which had been conducted objectively in order to prove or disprove the researcher’s hypotheses.

1.6.1. The research purposes

The primary topic running through this research is enhancing and achieving objectivity and reliability in assessing the EFL learners’ essays. Because to err is human, the researcher assumption is that human assessment to the learners’ written production may be not reliable even when such an assessment is built upon a careful specification of an analytical assessing device. Consequently, the research thinks that the intervention of the machine i.e. the computer has become more than just a must to increase the consistency and reliability of the assessment. The computer may do things much more effectively than the human does because of so many reasons chiefly its non-subjectivity and the extended data-base it may have. Thus, in order to increase the reliability of the assessing procedure, the research suggests the creation of some computing programmes built upon a specific analytic assessing architecture.
or scheme. The use of computers to assess writing is not a recent issue. In the United States of America, for example, the history of using computers in assessing the students’ written production goes back to 1960’s. It is an indispensable condition to take an over-restrictive view of what it is supposed to be tested (validity). Reliability does not guarantee validity, but it is an essential precondition. As a general rule increased reliability will always increase validity.

All in all, such a study aims at achieving reliability-validity tension (Davies, 1978), which are the most significant features of the psychometric tests as opposed to those of pre-scientific days which are highly subjective and unreliable (Heaton, 1991: 16). Its main and fundamental objective is to come to an objective and reliable assessing procedure. Such objectivity is determined by consistency in scoring. Accordingly, the written production mainly the essay is supposed to be given a reliable assessment whenever it is assessed and whoever assesses it.

1.6.2. The research problematic

Assessing any essay is subjective in the sense that its merit has to be evaluated or judged by the examiner as a test writer and classroom teacher. Both the learners and the teacher view the interpretation of the answer and the way it must be undertaken from an extremely subjective opinion. The learners’ responses to the subject-matter are entirely subjective in the sense that each one approaches the essay topic from a distinct subjective perspective and adopts a highly different strategy. On the whole, each learner expresses his own opinion freely using his own verbal skills, interprets information in a way he desires and approves an absolutely personal organisation of his ideas.

The teacher, however, finds himself in face-to-face with different types of answers. Each answer differs from the others in terms of content, style, and organisation. His assessment to the appropriateness of the learner’s written response is influenced by his opinion of its
content, structure, style or even whether he agrees with what the learner has uttered. In such circumstances, the judgement of the worth of the written essay is tremendously tricky. The teacher cannot avoid subjective interpretation of its content or of its writing process. So, he may get in an embarrassing situation where he would be completely unable to provide a conclusive reason to the marks he gives. His decision about the kind of mark to be given to each essay seems inconsistent. The same paper, as a fact, is likely to be assessed and scored differently in different occasions. Moreover, it is probable that different but equally teachers assessing one essay can assign dissimilar marks. Their judgments about the correctness of the essay are highly subjective in the sense that each teacher judges the quality of the essay according to his knowledge, and the criteria he relies on in his writing assessing procedure.

Consequently, three major aspects may result from the assessment of an essay. First, the essays’ scores tend to be unreliable. Second, the students would be aware about neither their level of proficiency in writing and nor their own lacunae. Third, the writing teacher would ignore everything about his learners’ deficits in writing thus he would be unable to design an appropriate effective feedback to enhance the learners’ writing skills. Such factors reflect the real existence of subjectivity in assessing an essay and which in turn makes no doubt that such subjectivity is still a problematic.

As opposed to subjective test such as the essay-question, objectivity is quite high in the so-called objective tests such as cloze-test items, or true-false answers. The relevance of such kinds of tests is that there is no subjective judgment on the part of the teacher, whose assessing procedure seems consistent. Besides, the same essay could be given the same mark even if it is assessed by different teachers. There is always an objective judgment of the response. The answer is always either right or wrong. No personal information is needed from the part of the learner, who is supposed to complete the assignment by using the language of
the teacher not his own. Accordingly, the assessing procedure is carried out as objectively as possible. No space is left to subjective judgments and personal evaluation.

All things considered, the present humble research focuses on the existence of many factors which may shape the assessing procedure of an EFL learner’s written production mainly the essay. Such factors may increase unreliability of the assessment procedure. They may also deprive both the teacher and the learner from any impressions about the evolution the writing skills. Therefore, assessment may lose the reasons for which it was created. Those factors might be summarised in four tips: the subjective judgement of the teacher, the unplanned assessing procedure, and the vagueness of the essay question.

1.6.3. The research questions

The proposed study seeks to answer three queries which are formulated on the basis of a review of issues and researcher primary investigations and observations.

1-Is the teacher’s assessment to the EFL learners’ essays reliable enough to achieve effectiveness?

2-How can an objective and reliable assessment to writing be achieved?

3-Can the computer be an alternative procedure to enhance the reliability of the essay assessment?

The researcher believes strongly that the writing assessment undertaken by the teachers (at least the teachers working on the same ground namely the University of Djillali Liabes) is not achieving its aims. The great numbers of students at the English department, the inappropriate assessing procedure, the ambiguity of the writing activities are few reasons among others that should be reviewed deeply to provide satisfying answers to the previous questions.

1.6.4. The research hypotheses
After three years (or may be five years for the master students) of being exposed to English through learning the four language skills mainly writing, few students are those who could develop an effective writing style. Such a fact is easily noticeable in the writing of the learners’ either Licence or Master theses. The researcher, therefore, doubts the effectiveness of the teachers’ feedback. And since the teachers’ feedback is to a great extent dependent upon the assessment procedure, we can deduce the following hypotheses:

- The teacher’s assessment to his EFL learners’ writing is not reliable and objective so as to achieve effectiveness.
- To achieve reliability and objectivity in assessing the EFL learners’ writing, the teacher may map an analytic assessing formula.
- Such a designed analytic assessing formula may function more effectively if converted into a computer programme or software.

An objective and reliable assessment provides an effective feedback, which in its turn affects certainly the mode the students write. The computer could be of a great help to facilitate the issue of assessment and make the burden less heavy.

1.7. The research population

The present enquiry engages two subjects: main subject or the third year English LMD students, and secondary subject or the second year English LMD students. Both subjects are from the English Department of the University of Djillali Liabes, Sidi-bel-Abbes. The purpose of including these different subjects is the main concern of this study. They both contributed to highlight the problematic which is still a matter of investigation. Second year students were submitted to an objective test such as cloze test items. Third year students, on the other hand, were submitted to an essay test. The difference between the testing procedures brought to light clear results of the two scoring procedures.

1.7.1. Main Population / Third year students
The study involved an experimental group of 92 third year EFL students who had been taught Linguistics at the department of English, University Djillali Liabes, in Sidi-bel-Abbes, during the academic year 2013-2014. They were divided into four groups: A and B. The group A consisted of forty-seven students. The group B consisted of forty-five students. Each group received one an hour and half course in Linguistics each Sunday. The group A started at 8:30 o’clock a.m., until 10:00 a.m. The group B began just after the group A and finished at 11:30 a.m.

These subjects contributed greatly in determining the real existence of subjectivity in scoring an essay and the ineffectiveness of the teachers’ assessment. The first semester exam was the starting point which at some extent brought to light a very serious problematic. It was impossible to carry on the scoring procedure in an objective way, because of the lack of experience in the field of testing. The students complained about the marks they had been given, and it was difficult, first, to control the assessing procedure, and to convince the students of why and how they had obtained such marks.

1.7.2. The main population linguistics syllabus

The following programme was tailored by the teacher according to the guidelines provided by the Ministry of Higher Education and Scientific Research. The curriculum was a kind of an attempt to combine a review the linguistics lectures that the students dealt with previously during their first and second years of the study along with some initiations to sociolinguistics. The following programme was taught to the EFL learners at the University of Djillali Liabes, Sidi-bel-abbes during the period from the mid of October 2013 to the beginning of May 2014. One hour and an half was allocated to each session once a week. Because of the inadequacy of the time allocated to teaching Linguistics the programme has never reached its end. Most of the time, teachers tackle the most important points in the syllabus in order to make a better use of the allocated time. It is worth mentioning that such
trouble concerns the entire subjects not only linguistics. The issue has been raised by teachers several times, yet no response from the part of the administration is noticed.

<table>
<thead>
<tr>
<th>First semester</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Weeks</strong></td>
</tr>
<tr>
<td><strong>Months</strong></td>
</tr>
<tr>
<td>September</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>October</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>November</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>December</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>January</td>
</tr>
</tbody>
</table>

| Winter holidays | First Semester Examination |

Table 1.3. Third year teacher adapted Linguistics syllabus (1st Semester)
Table 1.4. Third year teacher adapted Linguistics syllabus (2nd Semester)

1.7.3. Secondary Population /Second year students

The study also involved a control group of just 60 second year EFL students who were taught Linguistics, at the department of English, University Djillali Liabes, Sidi-bel-Abbes, during the academic year 2013/2014. The second year students were more than 160, but the researcher chose randomly 60 students. They were divided into two groups: A and B. They had been taught by two different teachers who were in charge of the same curriculum. Each teacher had his own programme of teaching. The testing procedure differed from one teacher to the other, but the objective remained the same.

The two groups A and B received one an hour and half as Linguistics course each Tuesday morning. The group A consisted of 32 students. It started at 9.30 a.m. until 11.00 a.m. The group B consisted of 28 students. It started at 11.00 a.m. until 12.30 a.m. Those Students were
submitted to cloze test assessments rather than to essay question assessments. The researcher tried to decipher the ease of assessing a cloze test and the reliability of its scores.

Cloze test or gap-filling is an objective test which impels the learner to write the accurate item. The possibility to react perfectly depends to some extent on the students’ knowledge of what they had learned. They were assigned to fill in the blanks with the right words. Such a task permitted to achieve easy assessing procedure. The students are allowed to an auto-evaluation. They can know by themselves how and why they had obtained such marks. Therefore they would be able to recognise their strengths and their weaknesses about the subject matter.

The inclusion of cloze test question contributed greatly to make a clear distinction between the so-called subjective and objective tests. Such a contribution can be seen in three main ways: first, the assessing procedure had been easily conducted; second, it was quite possible to control the reliability of the assessing procedure, and third, the students were able to predict which mark they would be given.

1.7.4. The secondary population linguistics syllabus

The following curriculum was adapted by the linguistics teachers according to the guiding principles supplied by the Ministry of Higher Education and Scientific Research. The syllabus was an attempt to introduce linguistics with its different subfields. The programme as it noticeable is divided into three part. The first part deals with linguistic as a scientific study of the language with a focus on the characteristics of scientific approach. It also sheds light on human and animal languages through contrasting both. The second part is reserved to the language and its different levels namely its pronunciation, grammar and semantics. The last part of the syllabus is about the branches of linguistics mainly the micro- and the macro-linguistics. The following programme was taught to the EFL learners at the University of Djillali Liabes, Sidi-bel-abbes during the period from the mid of October 2013 to the
beginning of May 2014. One hour and an half is allocated to each session once a week. Because of the insufficiency of the time allocated to teaching Linguistics the plan has never attained its conclusion.

<table>
<thead>
<tr>
<th>First semester</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Weeks</strong></td>
</tr>
<tr>
<td><strong>September</strong></td>
</tr>
<tr>
<td><strong>October</strong></td>
</tr>
<tr>
<td><strong>November</strong></td>
</tr>
<tr>
<td><strong>December</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>January</strong></td>
</tr>
</tbody>
</table>

Table 1.5. Second year teacher adapted Linguistics syllabus (1st Semester)
| February | II- Levels of language: speech and writing | 2-Level of speech sound.  
  a-Phonetics & Phonology  
  The sounds of language | b- Phonetics & Phonology  
  The sounds of language | b- Phonetics & Phonology  
  The sounds of language |
|----------|------------------------------------------|-------------------------------------------------|-------------------------------------------------|-------------------------------------------------|
| March    | B- Grammar:  
  1-Writing as a secondary system of communication. | 2-Components of writing:  
  a-Morphology  
  The structure of words | | Spring Holidays |
| April    | a-Morphology  
  The structure of words | b-Syntax  
  The structure of sentences | b-Syntax  
  The structure of sentences | C- Semantics  
  The study of meaning  
  Types of meaning |
| May      | III-Branches of Linguistics  
  A-Micro-linguistics  
  1-Structuralism  
  2-Generativism | B-Macro-linguistics  
  1-Sociolinguistics  
  2-Psycholinguistics | | Second Semester Examination |

Table 1.6. Second year teacher adapted Linguistics syllabus (2nd Semester)

Teachers do not always agree about the same syllabus and so are the linguistics teachers. At the university, there is a flagrant lack of coordination between teachers.

1.8. Data collection devices

The broad approach used in this study was to collect data using various procedures. Such procedure has included two tools: two written tests for both populations mainly the second
and third year EFL learners at the University of Djillali Liabes, Sidi-bel-Abbes and a teachers’ interview (see appendix 13).

The present research took into consideration the first semester exam session that the students were solicited to respond to. Administratively speaking, the aim of the examination, as usual, is to provide a final mark for the whole semester work, upon which a decision is to be made about either the success or failure of the learner to join the upper level. Pedagogically speaking, on the other hand, such an exam is designed to evaluate and assess students’ mastery of the courses taught during the whole semester. The students were invited in advance to review the whole courses that they had been taught.

1.9. Secondary population’s written test

Second year students were given a cloze test as a part of their exam process on Tuesday, January 28th, 2014. Such a test format (see Appendix 1) had been designed in order to evaluate students’ abilities to decide on the correct items. The test question was derived from the course sessions, more precisely from ‘Introduction to linguistics’ lecture to 'Human language' one. Such lectures had been taught during a span of time of 18 hours, i.e., from Sunday, October 20th, 2013 to Wednesday, January 15th, 2014. The students were assigned to fill in the gaps with the right words. The test consisted of forty blanks that the students had to complete. Each blank is allotted 1/2 a mark. The whole answer is out of 20. The students had to write down correctly the appropriate word. They might lose 0, 25 for each error. The instructions were written in the exam questions so that the students could evaluate their own answers. The classroom teacher or the researcher assessed and scored 60 exam papers in less than three days. Each exam paper submitted to double correction in order to validate the assessing procedure. Such an assessing process had been undertaken in three different occasions. The possibility to encounter difficulties was very low because of the strict control
of the assessing procedure. The answers to the cloze test question were either true or false. No additional information from the part of the students were needed.

1.9.1. The students' contributions to the assessment procedure

Second year students were invited to assign marks to their own answers. They were given a corrected examination sample to be guide in their assessment to their own written productions (see Appendix 3). The teacher's objectives were first, to provide the students with an opportunity to judge the quality of their own works or to auto-evaluate, and second to check the validity and reliability of the marks he had allocated to the same exam papers.

Second year students were divided into two groups as mentioned previously. Both groups included about 60 students. However, it was not possible to rely on the contributions of such number of students. Just a limited number could be available. The selection of samples had been applied for two purposes: the selection of two exam-papers and the selection of three students who were supposed to mark and assess such writing samples (see Figure 1.3.)

1.9.2. The selection of the written-papers

A great care had been exercised in selecting two exam-papers from a total number of the whole population. The classroom teacher relied on the students’ contributions in order to select one exam-paper from each group. Thus, in order to obtain a valid selection, a system called ‘the sampling frame’ had been adopted. This system had been applied in both groups, in different occasions. In each group the same techniques were used. Each student was allocated a number. The total number of students indicated the number of paper slips, each bearing a number from one to the total size of the population of the group. The set of paper slips were put into a box. One student was given the task to represent his peers. He was asked to mix up the slips thoroughly, and withdraw one paper. This selected paper slip would represent a number. The number, in turn, would represent the exam-paper, and the exam-paper would represent the group in which the sampling frame had been used. The two exam-
papers, therefore, were given the letters ‘A’ and ‘B’. The letters represent both the students and their groups. The letter ‘A’, for instance, represents the student ‘A’, and the group ‘A’. The letter ‘B’, on the other hand, represents the student ‘B’, and the group ‘B’.

1.9.3. The selection of students as assessors

The classroom teacher had used the same sampling frame procedure in order to select assessors. The assessors were students who were given the task to mark their peers’ exam-papers. However, they had not to belong to the same group of exam-paper to be assessed. That is to say, they ignored everything about to whom the exam-paper belonged. The same sampling frame had been used to select two exam-papers from both groups. These two exam papers represented two students who were allocated the letters ‘A’ and ‘B’. These two students were not to be known by their peers’ assessors. The classroom teacher wanted three scorers to be selected from the total mass of the population of each group. The selection had taken the same sampling frame techniques applied in selecting the two exam-papers. In each group, the students were assigned to select the desired numbers of students. The selected students were given the numbers from one to six. In group ‘A’ the three students were given the numbers 1, 2 and 3. In group ‘B’, on the other hand, the students were given the numbers 4, 5, and 6. Every student hoped to participate to the process of assessment. They were all motivated and excited to contribute to the operation. They admitted that this was the first time that a teacher provided them with such an opportunity of recognizing and assessing their own written paper deficits. Pedagogically speaking, every time, the learners are involved, the effectiveness of the teaching learning process increases and vice versa. The following figure illustrates the methods used by the researcher to select the written tests samples and the assessor students
1.9.4. The assessing procedure:

Before the students were given the instructions, the classroom teacher made photocopies of the two selected exam-papers. For both exam papers fourteen photocopies were made. Each exam-paper was to be submitted to seven different assessments, thus seven photocopies were needed. The classroom teacher hid the marks that he had previously given, and the students’ names. The students were asked to track the following procedure (see Figure 1.4).

1.9.4.1. The first procedure: Self-Assessment

The students ‘A’ and ‘B’ were assigned to give a final mark to their own works. Each student was given a photocopy of his own answer. He was kept in his immediate group. Both students were equipped with the exact cloze test’ answer sample (see Appendix 3). The objective behind such an attempt was to check to which extent the students’ assessment would comply with the teacher’s one i.e. to check if the students could give out at the same marks as those of the teacher.

1.9.4.2. The second procedure: Assessment exchange
The two students ‘A’ and ‘B’ were asked to assess and score each other’s exam paper. The student ‘A’ was asked to assess and score the exam paper ‘B’, of the student ‘B’. And the student ‘B’ was asked to assess and score the exam paper ‘A’ of the student ‘A’. The classroom teacher made one photocopy for each exam paper. The procedure consisted of keeping each student in his own group. Both students had to ignore each other. The main objectives in doing so, were to make students judge each other’s work, and to see if they could arrive at the same results that had been assigned by their teacher.

1.9.4.3. The third procedure: Multiple-assessment

The classroom teacher made six photocopies for each exam paper. He gave them to the students assessors in order to give marks to the exam papers ‘A’ and ‘B’. In group A, the students 1, 2 and 3 were given six photocopies of the exam papers ‘A’ and ‘B’. In group B, the students 4, 5 and 6 were given also six photocopies of the exam paper ‘B’ and ‘A’. The students as assessors in both groups ignored each others, and they also ignored to whom the exam papers belonged. Each of them was assigned to assess and give a final mark to their peer’s cloze test answer. The instructions had been done in a strict confidence. Each student in each group was asked to hide his marks. No one could see what the other or others were doing. The objective of such a multiple-scoring was to see if the six students could obtain the results as those obtained by the teacher and their peers ‘A’ and ‘B’. The following figure sheds light on the process that the students as assessors have gone through in assessing their peers’ written papers. The figure summarises the architecture of the students assessing procedure as designed by the researcher. The letters ‘A’ and ‘B’ represent the selected written essays samples whereas the numbers ‘1’, ‘2’, ‘3’, ‘4’, ‘5’, and ‘6’ refer to the chosen students who were supposed to assess their classmates essays.

![The assessing Procedure](image-url)
1.10. Primary population written test

Third year students had been given an essay-type test as part of the examination process, on Sunday, January 26th, 2014. It is a content/subject examination (Carter; 1973: 22) in which the students are asked to discuss, i.e., to present the different aspects of the suggested test question, in a limited time of about an hour and half. The main advantage of an essay is that it would tell the examiner or classroom teacher a lot about the students’ ability in a particular skill (Nolasco and Arthur, 1983:17). In addition, the essay-based testing is thought to encourage a better conceptual understanding of the material on the part of the students and to reflect a deeper, more useful level of knowledge and application by students. Thus, grading and criticizing written products is important not only as an assessment method, but also as a feedback device to help students better learn both content and the skills of thinking and writing. Such performance assessments are seen as differing from tests such as multiple-choice testing or cloze tests in that students create responses rather than selecting them, and moreover performance assessments emphasize problem solving and other higher-level integrative cognitive skills.

The open-ended question (see Appendix 2 / essay question) assigned the students to elicit completely subjective responses. They were asked to use their own words and writing processes. The students had been previously taught how to respond to such a question with its key word ‘discuss’. The question was derived from course sessions, more precisely
‘Introduction to Sociolinguistics’ which represents the first lectures that the students dealt with. Those lectures had taken place during the period from Sunday, October 20th, 2013 to Sunday, December 8th, 2013. The double objectives of the teacher in proposing such an exam were first, to determine the extent to which the students had mastered the courses taught in the formal classroom sessions, and second to assess the students’ knowledge of the whole language. In such a situation, the teacher adopted two possible tests: attainment and proficiency tests. These tests had been integrated as interrelated parts of the scoring procedures. They could enable the classroom teacher to determine a student’s ability to put facts into perspectives, to generalise from the data he had mastered, to draw subjective conclusions from the content of his work, and to measure how well he was able to communicate in writing.

1.10.1. The first tentative assessment: Holistic assessment

The first attempt to assess the EFL learners’ written production was a kind of impressionistic assessment. The researcher assessed his own learners’ essays. He also gave the investigation another dimension by inviting his colleges to assess some samples of the students’ written pieces. The classroom teacher or the researcher wished to invite at least ten other teachers to participate to the process of assessment to have reliable and valid results, unfortunately but it was really impossible to do so.

1.10.1.1. The researcher’s assessments

The classroom teacher or the researcher spent 15 days to correct 92 written papers. It was difficult to state very precisely which criteria were to be tested. Each exam paper was to be judged according to the teacher’s feeling and prejudice. The scoring procedure adopted was quite holistic. Credit was not given to specific criteria. The classroom teacher encountered serious problems in his assessing procedure. It was impossible to make a conclusive decision about the quality of the answers, and the final mark to be given to each exam written piece.
Even if there had been clear differences between the answers, it was entirely impossible to say why such an exam paper had been given such a mark. Therefore, the researcher found himself assessing each exam paper three times in many different occasions. The table shows the time consumed by the researcher in his way towards achieving a reliable assessment to his learners’ written production.

<table>
<thead>
<tr>
<th>The researcher’s assessments</th>
<th>Number of the exam papers</th>
<th>Assessing time</th>
</tr>
</thead>
<tbody>
<tr>
<td>The first assessment</td>
<td>92 exam papers</td>
<td>3 days</td>
</tr>
<tr>
<td>The second assessment</td>
<td>57 exam papers</td>
<td>10 days</td>
</tr>
<tr>
<td></td>
<td>35 exam papers</td>
<td></td>
</tr>
<tr>
<td>The third assessment</td>
<td>46 exam papers</td>
<td>1 day</td>
</tr>
<tr>
<td></td>
<td>46 exam papers</td>
<td>1 day</td>
</tr>
</tbody>
</table>

Table 1.7. The researcher’s assessments

The assessing procedure took three steps. In the first step the classroom teacher corrected all the exam papers in three days. His judgements about the quality of the students’ answers were not conclusive. He could not prove why he gave such marks. He could not decide upon the quality of the students’ written pieces. The second attempt was undertaken two days after the first scoring. Such a scoring procedure was divided into two steps. It took the classroom teacher 10 days to correct again all the exam papers. In the first step, it was possible to correct 57 exam papers in 6 days. Two days after, the remaining exam papers were corrected in 4 days with great efforts. Although there had been a double assessing procedure, the probability to reach conclusive results was still a matter of doubt. However, in order clear up such a doubt a third and a last scoring procedure had been set forth. It was impossible to correct all the exam papers in one day. The classroom teacher split up the number of the exam papers into two parts. Each part consisted of 46 exam papers. The scoring was undertaken in two days.
Each day was devoted to each part, with a rest of three days between the two steps of scorings.

Even if the classroom teacher had tried in different occasions to make his scoring sound fair and reliable, the problem of uncertainty was still pertaining his scoring procedure. However, in order to find a solution to such an inquiry he sought help of some teachers who taught the same module.

1.10.1.2. The Teachers’ contributions: Multiple assessments

The problem of unreliability in scoring caused by the impressionistic method had impelled the classroom teacher to enlist services of three of his colleagues who had worked on the same ground, i.e., the teaching of Linguistics and Sociolinguistics. At the beginning, it was not possible to find someone who was ready to devote time and efforts in order to correct the exam papers of other students and not his or hers. The few teachers that accepted to cooperate suggested correcting no more than two exam papers.

Before providing the teachers with the exam papers in order to be corrected again, the classroom teacher selected two exam papers from two different groups namely ‘A’ and ‘B’. In each group one exam paper had been selected as a multiple scoring sample. In doing so, the scoring procedure might carried in a more objective way. It was possible to enlist the contributions of the students who were not convinced of the marks they had been given. The selection of the two exam papers had taken the same method that had been applied with the second year students. The same sampling frame had been adopted in each group. The selected two exam papers were referred as (A) and (B) in reference to their groups. The exam paper (A) refers to group (A) and the exam paper (B) refers to group (B). They had been handled to the other teachers a day after The classroom teacher omitted the students’ names and the marks that he had previously allotted before making photocopies. He made three photocopies for each exam papers. Each teacher had been given two photocopies of the exam papers (A)
and (B). They were asked to use the essay-answer (Appendix 4), before starting reading the students’ exam answers. Such an answer is presented as the main points of discussion that the students had to tackle. The teachers’ contributions were an attempt to prove or disprove the problem of the assessment unreliability. Such a procedure took about a week. The table shows the time consumed by the teachers who were solicited to assess the learners’ written production.

<table>
<thead>
<tr>
<th>Assessors</th>
<th>Assessing time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher 1</td>
<td>Four days</td>
</tr>
<tr>
<td>Teacher 2</td>
<td>One day</td>
</tr>
<tr>
<td>Teacher 3</td>
<td>One week</td>
</tr>
</tbody>
</table>

Table 1.8. Teachers’ assessing time

A week later, the researcher collected the exam papers from each teacher. The marks allotted to the given exam papers differed to a great extent from those the researcher had given. Furthermore, there had been a disagreement over the marks given by the teachers. The teachers’ judgements differed from each other. Such a disagreement impelled the classroom teacher to beg his colleagues to answer the proposed questionnaire.

1.10.2. The second tentative assessment: Analytic assessment

In order to avoid the problems that had been encountered with the impressionistic assessing, an analytic method had been adopted. This method depends chiefly on a carefully drawn analytical scoring plan (see Appendix 9). The main objective of such a scoring device was to achieve a more reliable assessing procedure. The analytic scoring scheme consists of six selected criteria which are of two kinds. The first includes all those criteria that has been underscored by communication approaches to language teaching and includes the content (the substance of writing), the organisation (cohesion and coherence of ideas), and style (the accurate clear use of words and sentences). The second includes all those criteria that focus on
the accuracy of language, such as grammar (the ability to write accurate sentences),
mechanics (the effective mastery of the graphic conventions) and diction (the correct choice
and use of words).

The six selected criteria had been carefully drawn from the first examination which had
been designed to third year students. It was not possible to pay more attention to the same
criteria all together, because of the impressionistic assessment that had been adopted. There
had been a tendency to rely on the gist of the content, grammar, organisation, legibility and
neatness of the handwriting, style, mechanics and diction. These eight criteria were not
consistently used. Their usage changed from exam-paper to another. Some exam papers were
marked according to their contents, grammar, and handwriting. Some others were marked on
the basis of the neatness of the exam papers, the style, mechanics, diction and organisation.
However, in order to avoid inconsistency in using these criteria, six of them were adopted as
standard basis for the scoring procedure. More emphasis was given to the content,
organisation, style, grammar, mechanics and diction. The legibility and neatness of the essay-
answer were not taken into account because of their slight weight within the process of
writing.

1.10.2.1. The researcher’s assessment: The threefold assessment

In order to avoid scores unreliability the classroom teacher adopted the Analytic Method
which is mainly dependent upon a careful specification of a psychometric assessing
procedure. The analytic device that goes with the applied method was constructed in two
days. The classroom teacher used the same two exam papers that he had given to his
colleagues. The probability to achieve an objective and reliable assessment was still a matter
of doubt. However, in order to clear up such a doubt threefold assessment had to be applied.
The assessing procedure had gone through three steps. Each exam paper was corrected three
times as it is shown in the following table which displays the dates of the assessments undertaken by the researchers.

<table>
<thead>
<tr>
<th>The threefold assessment</th>
<th>Dates of assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>The first assessment</td>
<td>February 2\textsuperscript{nd}, 2014</td>
</tr>
<tr>
<td>The second assessment</td>
<td>February 5\textsuperscript{th}, 2014</td>
</tr>
<tr>
<td>The third assessment</td>
<td>February 8\textsuperscript{th}, 2014</td>
</tr>
</tbody>
</table>

**Table 1.9. The researcher’s threefold assessment**

The classroom teacher proceeded by dividing the scoring procedure into three steps. In each step he corrected the two exam papers at the same time. He also made a bridge of two days between the three steps. Such an attempt was done in purpose in order to avoid similar scoring judgements. The analytic scoring procedure adopted permitted him to follow step by step the scoring of each exam paper. The following figure presents the assessing procedure designed by the researcher.

![Diagram](image)

**Figure 1.5. The researcher’s adapted assessing procedure**

The analytic assessing procedure consisted of dividing the scoring procedure into six steps according to the number of the assessing criteria. The researcher started by judging the quality of the answer according to his knowledge of the subject-matter. The second step is mainly dependent upon the first one. The correctness of the content is to be regarded as a starting point to the other criteria. The relevance of the essay answer is determined by whether the
student has come to understand the question or not. If the answer is regarded off-topic, all the other criteria are deemed irrelevant.

The teacher, thus, carried on his assessing procedure on the basis of the relevance of the ideas expressed. The other criteria are also judged step by step. Each exam paper was assessed according to the proposed method. The three steps of assessing were based on a psychometric basis. In each step, a counting procedure was used. Each assessing criterion was first judged and allotted a specific mark. At the end of each assessing procedure a marking grid had been used. This marking grid (see Appendices 11 and 12) consists of the all the assessing criteria with the allotted marks.

1.10.2.2. The teachers’ assessments

The classroom teacher enlisted services of the same three teachers who had previously contributed to the first scoring attempt. They have been equipped with all the assessing tools that the classroom teacher had used. They were given two photocopies each of the same exam papers that they had previously corrected. They were also begged to use the same scoring steps. In a sense, they were asked to follow the same steps that the classroom teacher had adopted. The teacher’s objective in doing so was to see if the analytic assessing method could achieve the same scores as he had already allotted to the intended exam papers. The time consumed by the teachers to assess the written papers is displayed in the following table

<table>
<thead>
<tr>
<th>The teachers’ assessment</th>
<th>Dates of assessments</th>
</tr>
</thead>
<tbody>
<tr>
<td>The first teacher</td>
<td>February 10th, 2014 to February 12th, 2014.</td>
</tr>
<tr>
<td>The second teacher</td>
<td>February 10th, 2014 to February 13th, 2014</td>
</tr>
<tr>
<td>The third teacher</td>
<td>February 11th, 2014 to February 12th, 2014</td>
</tr>
</tbody>
</table>

Table1.10. The teachers’ assessment through the analytic method

The teachers were begged to render the corrected exam papers in one week. On February 16th, 2014 the exam papers were collected. The problem in assessing the essay was not totally
reduced. Scores unreliability was still a matter of investigation. In chapter four an attempt was made to provide some recommendations in order to achieve a formal and reliable assessment of an essay.

1.11. The teachers’ Interview

The three teachers who participated to the assessing procedure have been begged to answer as freely as possible the intended questions proposed in the interview (see Appendix 13). The interview had been set forth in a week. The interview was an attempt to decipher many sides of the issue of assessing essays. The teachers were asked about their attitudes and preferences towards the writing assessment process in particular.

The interview consists of six questions. The first question was about the teachers’ use of the essay as an effective tool to assess their students’ writing. The second question was about their preferences about the methods they used to assess their learners’ written productions namely the error-count method, the multiple assessing method, the analytic assessing method or the holistic method. The third and fourth enquiries were of a paramount significance for the study since they intended to check up the teachers’ outlook and rates to each of the following essay’s scale: excellent, very good, good, high average, low average, weak, very weak and off-topic. The fifth question tried to determine which components or areas the teachers usually consider in their assessment to a written essay. The last question was an attempt to check if the teachers’ assessment was always effective which may result in an effective written feedback supplied by the teachers themselves. The teachers were given a total freedom to answer the questions. All in all, the interview was meant to decipher the issue of assessment unreliability and its ineffectiveness in generating a useful and efficient feedback that could enhance the EFL learners’ writing aptitudes and competences.

1.12. Methodology
The foregoing research was based upon considerable data, which had been collected through the use of different tools such as a corpus of both second and third year EFL learners’ first semester examination and a teachers’ interview. As far as the second year students’ examination is concerned, the focus was, firstly, on the assessments provided by the classroom teachers, and secondly on the assessment provided by the students themselves to who were invited to assess their classmates’ written papers. The contribution of the learners to the assessing procedure is of a paramount significance. On the other hand, third year students’ examination enriched such a study with its twofold attempt of assessment. The data, thus, had been collected from both assessing methods of the researcher and the contributions of the teachers in both attempts. The following table recapitulates the researcher methodology use to test the hypotheses.

<table>
<thead>
<tr>
<th>Research tools</th>
<th>Students contributions</th>
<th>Research procedure</th>
</tr>
</thead>
<tbody>
<tr>
<td>2nd year Written-test</td>
<td>The researcher’s assessment</td>
<td>First assessment attempt</td>
</tr>
<tr>
<td>(cloze test)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3rd year Written-test</td>
<td>The researcher’s assessment</td>
<td>First Holistic assessment attempt</td>
</tr>
<tr>
<td>(essay)</td>
<td>The other teachers’ contributions</td>
<td>First Analytic assessment attempt</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The interview</td>
<td>The teachers who contributed to the different assessment procedures of the students’ essays were invited to answer some questions set by the researcher to clarify many ambiguous issues</td>
<td></td>
</tr>
</tbody>
</table>

Table 1.11. The researcher methodology
The data was collected over a span of time of almost four months. The classroom teacher devoted three months as experimental learning with the third year students, and less than twenty days with the second year students who were regarded as secondary subjects of this study. The problem that had to be resolved was very serious, and for this reason the probability to come to the wanted results was not completely achieved. The classroom spent more a year and half later in order to find acceptable solutions to the frequent problem that covers the assessing and scoring of essays. Such solutions are still in theory.

1.13. Conclusion

The inclusion of an essay-type test has been formally as a subjective phenomenon. Such subjectivity caused a trouble which is recognised in assessment unreliability and which is the main result of the holistic method. The subjective judgements of the teachers can show a tolerable interval but can never built up a reliable assessment. The classroom teacher’s judgements even if they seem right are inconsistent. Marks change their weight from one occasion to another, and from paper to paper. The teacher’s decisions on the marks he awards follow a continuous alteration. This is due to the fact that the scoring procedure he has adopted is mainly influenced by his feelings and opinions. The same exam paper was marked differently in different occasions.

The data showed that such an unreliability in scoring had been recognised also among other teachers who share the same exam paper. The three assessor teachers who had been invited to have an assessing word about the students’ responses increased unreliability in the sense that they had differed first, with each others, and second with the classroom teacher, in terms of the marks they had provided. Such disagreements were felt as being the main result of the different procedures that each of them had adopted.

However, in order to determine the effectiveness of the assessing procedure, the analytic method had been used. This method had been tried up in different occasions by the
contribution of different teachers who were assumed to give an answer to the problem of assessment unreliability. Such an answer can prove or disprove the objective of the study. Is it possible to achieve an objective and reliable assessment of an essay? Or rather can we reduce subjective judgements and reach a more reliable and formal assessment?
Chapter Two

Enhancing the essay assessment reliability
Chapter Two  Enhancing the Essay Assessment Reliability

2.1. Introduction  67

2.2. Testing the Written Production  68
   2.2.1. Objective Testing  68
   2.2.2. Subjective Testing “The Essay”  69

2.3. Subjective Testing Specificities  70
   2.3.1. Scoring Unreliability  71
   2.3.2. Validity Deficiency  72
   2.3.3. Interpretation Difficulty  74
   2.3.4. Preparation Effortlessness  76
   2.3.5. Administration Simplicity  76
   2.3.6. Scoring Difficulty  77

2.4. The Essay Assessment Approaches  79

2.5. The Holistic Approach  79
   2.5.1. The Holistic Approach Deficits  80
   2.5.2. The Holistic Assessment New Forms  81
      2.5.2.1. The Administration of Short Essays  81
      2.5.2.2. The Multiple Assessing Method  85
      2.5.2.3. The Error-Count Method  91

2.6. The Analytical Assessing Method  95
   2.6.1. The Nature of the Analytic Assessing Method  96
   2.6.2. The Analytic Assessing Procedure  96
   2.6.3. The Analytic Assessing Method Deficiency  97
   2.6.4. The Analytic Assessing Method Advantages  98
      2.6.4.1. Scores Reliability  99
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.6.4.2. Security</td>
<td>99</td>
</tr>
<tr>
<td>2.6.4. 3. Feedback</td>
<td>100</td>
</tr>
<tr>
<td>2.7. Feedback: An Overview</td>
<td>100</td>
</tr>
<tr>
<td>2.8. The Feedback Effectiveness and Ineffectiveness</td>
<td>104</td>
</tr>
<tr>
<td>2.9. The Feedback Levels</td>
<td>106</td>
</tr>
<tr>
<td>2.9.1. Feedback about the Task</td>
<td>107</td>
</tr>
<tr>
<td>2.9.2. Feedback about the Process</td>
<td>108</td>
</tr>
<tr>
<td>2.9.3. Feedback about Self-Regulation</td>
<td>109</td>
</tr>
<tr>
<td>2.9.4. Feedback about the Self</td>
<td>111</td>
</tr>
<tr>
<td>2.10. Writing Feedback</td>
<td>112</td>
</tr>
<tr>
<td>2.11. The Writing Feedback Forms</td>
<td>115</td>
</tr>
<tr>
<td>2.11.1. Direct Corrective Feedback</td>
<td>116</td>
</tr>
<tr>
<td>2.11.2. Indirect Corrective Feedback</td>
<td>117</td>
</tr>
<tr>
<td>2.11.3. Metalinguistic Corrective Feedback</td>
<td>119</td>
</tr>
<tr>
<td>2.11.4. Focused Versus Unfocused Feedback</td>
<td>121</td>
</tr>
<tr>
<td>2.12. The Teacher’s Roles in Delivering Feedback</td>
<td>122</td>
</tr>
<tr>
<td>2.13. Conclusion</td>
<td>123</td>
</tr>
</tbody>
</table>
2.1. Introduction

In any consideration of classroom written tests, a distinction must be drawn between the rather objective tests and subjective ones. Such tests are generally prepared, administered, and assessed by the same teacher. They are mainly used to assess the students’ achievements of the instructional objectives. Achievement tests, as a fact, are used to indicate group of individual progress toward the instructional objectives of a specific study. They measure the extent to which students have mastered the specific skills or body of information acquired in a formal learning situation. Examples of these tests are final examinations at the end of a course of study. They have a single cut off point: the examinee either ‘pass’ or ‘fail’ the test, and the degree of success or failure is deemed important to both the examinees and the examiner.

Objective tests are considered as formal classroom tests because they are analytical, and need no subjective judgements from the part of the teacher. The assessing procedure is highly reliable, and quantified in scientifically precise terms. On the other hand, subjective tests, such as essay-type tests, are holistic and impressionistic. They have been recognised as informal tests because they are supposed to have an ad hoc basis. The assessing procedure is often unreliable. The possibility to achieve assessing reliability is still a matter of investigation.

The search for an objective and reliable assessing procedure to the EFL learners’ writing has led some examiners to adopt different types of methods. The assessing of an essay-type test, thus, has witnessed the emergence of four different approaches: Short-Type Essays, Multiple-Scoring Method, Error-Count Method and the Analytic Scoring Method. These approaches differ from each other in terms of the scoring procedures, but they have a common objective- assessment reliability- that is extremely impossible to achieve with the single holistic scoring procedure. However, one would wonder if a formal and objective scoring
procedure is possible through the use of one of the previous mentioned approaches, or in other words: Is it possible to achieve scores and assessment reliability with one of these methods?

2.2. Testing the written production

Assessment is the most significant component of any effective teaching since it may allow the teacher to recognise the real level of proficiency of his learners. It also provides the teacher with an idea about the learners’ lacunae thus he would be able to design the appropriate remedial work trying to enhance the situation. Assessment is to continue during the whole process of teaching not only through specific and limited tests as it is the case with most teachers in our universities today. Most of the time, the assessment procedure is done two times a year throughout those famous examinations at the end of each semester. Examinations could be classified under two headings subjective and objective ones. Assessing the written language can be seen from two different types of tests: objective tests where the learner is supposed to provide a specific correct answer and subjective ones where the learner is supposed to write an essay in which he may discuss or analyse an issue. The main difference between these two tests is mainly recognised in the assessing procedure. The first procedure seems to objective whereas the second one tends to be more subjective.

2.2.1. Objective testing

Objective testing is always referred to as standardised assessment. The main common assessing devices are cloze tests, and multiple-choice items. They have been used as language testing techniques after the psychometric revolution of the 1930s. Standardised tests were developed into writing assessment in order to overcome a number of the weaknesses that covered the testing of essays. The most serious of these problems is scores unreliability. But evidence have shown that, even if reliability is neatly perfect in the case of multiple-choice items where the scorer seems very often like a machine, the essay test is likely to be a very important testing device. From a historical perspective, the widespread use of an essay as a
testing device probably grew out of the back to basics movement which emerged in response to charge that many of the educational systems lacked the fundamental academic skills of writing. The purpose of essay tests was to integrate educational tests more meaningfully into instructional process by emphasising the importance of communicative language testing as a remedy and as a substitute to the psychometric movement which occurred in the period from approximately the mid-1960s to the early 1980s. The psychometric approach with its standardised tests such as multiple-choice items, and gap-filling and its emphasis on the twin concepts validity and reliability emerged as a reaction against the traditional testing of an essay which was regarded as highly subjective and unreliable. However, by the emergence of the modern approach in language testing, some instructors tried to develop a new technique that should accompany it (Wilkins; 1979: 82). Clearly, the responsibility for ensuring a better and acceptable testing procedure depends to a larger extent on the appropriate selection of the measurement device, which may elicit the modern evaluation of the writing skill.

2.2.2. Subjective testing “the essay”

The birth of the psychometric and the integrative approaches, as a fact, has ensured two factors: objectivity and reliability, but evidence shows that these approaches ignored totally the real assessment of the writing skill. Their main purpose is to avoid the side effects which covered the measurement of such a type of task.

The development of the modern approach, however, has brought a new flavour to writing assessment. The essay test has given a formal status to language testing. It has been thought as being the most appropriate technique that can make student demonstrate their abilities in writing. Such a test emphasises the importance of language performance that is to say, testing the communicative aspects of language such as the content, style, organisation of ideas, and paragraphing, more than language competence which emphasises the testing mastery of language such as grammar, vocabulary and mechanics. It provides the examinees with an
opportunity to demonstrate their abilities to organise language material using their own words and ideas and to communicate. This advantage of an essay would tell a lot about the student’s ability in a particular skill (Nolasco and Arthur; 1983: 17). It also requires free responses on the part of the examinees rather than the selection of correct answers.

The examiner’s scoring procedure in mainly dependent upon the elaboration of two indispensable tests: attainment and proficiency test. The former is one which aims to measure how much a learner has learned of what he has been taught, and the latter aims to measure a learner’s knowledge of the whole language (Corder; 1985:369). The quality of the essay answer can be regarded from two sides: the answer to the question (the substance of writing), and language form. It may assure two competencies: performance and competence. The examiner’s objectives in constructing an essay-type question are: (1) to measure the examinee’s achievement of the instructional goals, that is to say, to measure his progress toward the instructional objectives of a specific study, and (2) to measure his specific strengths in each component of the essay (i.e., to measure his abilities in grammar, mechanics, organisation of the ideas, style, and so on). Such a measurement has a single cut off point: examinee either passes or fails the test. One single mark can determine the degree of success or failure.

2.3. Subjective testing specificities

In language assessment the term subjectivity refers to the holistic or impressionistic evaluation of the learners’ products (essay-answers to a single long question). Such an evaluation in mainly recognised as informal. An informal evaluation is the type of evaluation which is broad and global. Some studies have shown that the evaluation is considered to be subjective whenever it carries the following characteristics. Some of these characteristics such as scores unreliability cause a serious problem in language assessment mainly when it comes to assessing writing. Teachers will never recognise the extent to which their learners are
improving their writing skills. Moreover, they will never be able to design and provide some remedial works to make things better i.e. to enhance and equip the learners with sufficient tools to identify and to overcome their lacunae.

2.3.1. Scoring unreliability

Subjective assessment refers to the assessing of productive skills such as speaking and writing. An essay test is a knowledgeable device used in most common subjective examination. It is a type of examination in which the learner is asked to describe, discuss, enumerate, compare, state, evaluate, analyse, summarise, or criticise involve writing at specified length. Such a test allows the learner to compose his own relatively free and extended written responses to problems set by the teacher.

In foreign-language assessment the learner’s responses may consist of single paragraphs or may be full essays in which the learner is rated and assessed not only on his use of grammatical structures and lexicon of the target language but also on his coherent ideas and their organisation. Grades for such free-responses assessment may also consider the learner’s use of the graphic convention –spelling, punctuation, capitalisation, paragraphing and even handwriting.

Such a type of assessment is highly subjective since teachers usually and unconsciously introduce their personal opinions when assessing their learners’ written performance. According to Corder (1985: 360), our judgments on an essay or a précis are almost inevitably influenced by our opinions of its content or lexical structure, or in an extreme case, whether we agree with what the writer. Such judgements have a negative effect on the assessing procedure. It is quite possible that the same exam-paper will be given different scores in different occasions. Assessment reliability in such a situation tends to low as opposed to objective assessment. The evaluation is impressionistic, broad and difficult to quantify.
2.3.2. **Validity deficiency**

The final source of the invalidity of the assessment is unreliability. If the assessment lacks reliability it will also lacks validity. In other words, the assessment measures something consistently. It cannot measure anything which seems invalid. We want our assessment to measure as accurately as possible what they set out to measure. But if for any reason we cannot place our confidence in the results we get, then we can scarcely regard the tests as valid (Corder, 1985:364). While in language testing and assessment, everything is subjectively valid; in technical language it is rather unreliable.

There are three main factors which show clearly the invalidity of our assessment. First, the assessment is regarded as invalid when it contains technical words that seem impossible to understand (face validity is lacking). Second, it is quite possible that the majority of the learners may find the assessing question difficult to respond to, or they may misinterpret it because the tasks they are asked to perform is uncommon to them. The teacher may have not prepared them in advance and they do not know how to tackle such type of tasks. The class session may be designed only for theoretical courses. The methodology of writing an essay may be dismissed and not taken into account by the classroom teacher who is supposed to measure students’ writing abilities. In such a situation the assessment is considered as invalid even if the question seems clear and has a direct relation with the courses they have been taught. Third, the assessment lacks validity when the teacher cannot know which criteria he has relied on in his rating procedure; when his judgements seem without any previously planned bases. This point is the most central one which shows clearly the close relationship between reliability and validity in assessing writing.

The writing assessment is considered informal and invalid when it is not followed by an analytical specification of an evaluating or scoring scheme. Such an analytical rating scheme can ensure both validity and reliability at the same time. In the absence of such a device, the
assessing procedure will be inevitably broad and inexact. As a fact, when we consider that a learner has done a good translation or a good essay we do not know very precisely what quality or qualities we have mastered and we are far from confident that our measure is a valid one (Corder; 1985:358). The specification of the analytical assessing scheme should be done in advance, before setting the assessment or the test, in order to ensure a correct and reliable evaluation of the supported criteria.

The assessment would seem valid if it deals with the essay as a kind of verbal communication (Yorkey, 1982:235). All the essay components such as content, form, grammar, vocabulary (word-choice), style, mechanics, must be taken into account. It is not to assume that a given score on language form necessarily allows conclusions to be drawn about the examinee's language performance. The whole of communicative event was considerably greater than the sum of its linguistic element (Clark, 1973:432).

In language assessment, the essay is used to reveal the quality of the learner’s language performance; how well he can communicate his ideas as precisely and correctly as possible; how well he can resolve the problem that he is exposed to. The assessment will lack its validity if the essay is designed just to reveal the learner’s competence in one area, grammar for example but frequently the student’s performance and success in accomplishing the task may be masked by errors and a tired marker who fails to make the necessary effort to respond to the writing as a means of communication (Heaton, 1991:149). If the assessment only focuses on grammar, it will not show the examiner how well students can write in English to express meaning. It is a truism to say that in assessing an essay we need to know not simply the student’s ability in writing correct English, but also how he can communicate his thoughts.

The assessment needs to engage first and foremost with the communicative purpose and overall coherence and organisation of the student’s output, not only with localized errors
which should be a secondary concern. Innovative and interesting, writing is presented as a problem-solving task which challenges, rather than defeats the students (Hamp-Lyons: 1987). The classroom teacher should set forth the objectives of the task that he wants his students to perform. Everything which may cause a handicap towards students' writing abilities should be clearly cleared up. The students should be well-informed of how to tackle the problem imposed and how their responses are to be scored. In such a situation, it is possible to say that the test is legally and formally acceptable.

Unfortunately, this purely a theoretical conception of how validity must be conceived. In subjective assessment everything is hidebound by personal bias. Therefore, whenever we encounter such deficiencies in language testing, we declare the assessment as being subjective and invalid. Validity goes hand in hand with objective assessments when reliability tends to be high. Invalidity, on the other hand, follows subjective assessments when unreliability covers the rating or scoring procedure. It is the classroom teacher who can validate his test or invalidate it. He is the first responsible for the selection of the test materials which can be constructed either subjectively or objectively. The validity of the assessment is highly dependent on the manner in which the instrument is employed. Improper administration can invalidate and impair the performance of individual examinees and decrease the efficiency of the assessing procedure.

2.3.3. Interpretation difficulty

Interpretation is by definition an explanation of something which is not easily understandable. It is a conclusive result which can be determined by time as true or false. Free-responses tests are classified as subjective assessment because their interpretations are highly difficult. It is readily apparent that the learners are allowed to express their answers in their own words in a relatively unstructured testing situation.
The interpretation regarding the level of ability or correctness of performance on the test may be subjective. Each learner is likely to approach the test and the task it requires from a slightly different, subjective perspective, and to adopt slightly different subjective strategies for completing those tasks. According to Bachman (1990:38), these differences among test takers further complicate the tasks of designing tests and interpreting test scores. Two interrelated factors are supposed to increase subjectively in assessing written responses. Such factors may also cause difficulties in interpreting test scores and essay test-responses. According to Bachman, responding to a subjective assessment is determined by the use of two writing devices: styles and strategies. Every learner or test taker is able to express his own opinion freely and interpret information in any way he wishes. In the same way, the teacher is able to evaluate the quality of his own opinion and interpretation as well as the organisation and logic of his opinion. Style and strategies are, then, common factors which the teacher finds as real obstacles in his interpretations.

Styles are those general characteristics of both intellectual functioning and personality type that especially pertain and differentiate anyone from someone else. Whereas Strategies are specific methods of approaching a problem or task, mode of operation for achieving a particular end, planned designs for controlling and manipulating information. They are contextualized battle plan (Brown, 1987) that might vary from a moment, day or a year to another.

People are not machines, even though they are supposed to have the same devices such as ‘language-processing device’ or ‘language-learning device’. The fact is, of course, that our performance in any task is strictly personal. The learners can never approach the stimulus from a unique viewpoint, even if they are familiar with it. They cannot evade supporting their own opinions, rather than that of the examiner.
The teacher, also, cannot evade supporting his own opinions rather than of the learners. His subjective interpretation of the examinees’ responses is not authoritative but seems somehow intuitive. One of the most important characteristics of intuition is its nonverbalizability. Persons are not able to give much verbal explanation of why they have made such particular decision or solution (Brown; 1987: 249). Intuition as a human factor involves a certain kind of risk-taking. In subjective tests, examiners must be willing to risk techniques, methods, or assessment that may produce a vague and biased impressionism. That is precisely the difficulty.

2.3.4. Preparation effortlessness

Assessing the written language through an essay is very economical in terms of skill and time that are required to prepare them. It is just too easy for the teacher or the examiner to take pen in hand and to turn out his question in few lines. Indeed, the assessment objectives can be used directly on course objectives, and the assessing content derived from specific course content. The classroom teacher has just few words to say in relation to what he has already taught his students. The assessing question often includes item words such as discuss, comment, or explain. Quite possible, the classroom teacher may use quotations in terms of questions with specific item words instead of his own words. Such a procedure may decrease the time allowed to select the appropriate test question, and relieve the busy classroom teacher from straining his eyes and time in thinking about the assessing exercise.

2.3.5. Administration simplicity

Open-ended assessing tests such as the essay questions are commonly recognised for their ease of preparation and administration. Such tests are easily administered because of their short formats. The test directions are usually given verbally in a short time. Sometimes, the teacher hands out the assessing question to ensure that all the learners receive a better wording
of the question. The directions are paraphrased verbally in order to have a better understanding of the task.

2.3.6. Scoring difficulty

In its extreme form, subjective assessments are the ones where the teacher’s test question is derived wholly from his intuition or personal opinion i.e. holistic assessing procedure. Judgements are rather impressionistic and difficult to quantify. They are rendered in rather global terms (Brown; 1987:249). In such tests (open-ended tests), the assessment is not easy to achieve since writing proficiency involves numerous traits that seem difficult to define.

The teacher makes personal judgements about the quality of the learners’ responses. His assessments seem to be based on fallible opinions, which are mainly affected by extra-factors such as fatigue, carelessness, prejudice, etc. Such factors influence both what is to be assessed and how the assessment should be carried out. Studies in such a field have shown that the holistic procedure has become the best known one associated with the writing assessment. Its application on the assessing ground has not been well embraced by some instructors. Such a holistic assessment is mainly dependent upon an ad hoc basis. Although, it seems much faster, it needs a more planned and structured rating scheme, which can specify in an objective sense test assessment. The holistic approach has been adopted to gain a general impression of the learner’s responses, but some experiments in the field of assessment have shown that such an approach can never achieve reliability.

The chief difficulties that have been encountered in using essays as an assessment device are: (1) eliciting the specific criteria that the teacher particularly wishes to assess; (2) finding a way to evaluate these free-responses reliably and economically, and (3) making learners recognise their lacunae and their level of proficiency.

It is almost a truism to say that each skill is uniquely difficult, but assessing through essays in itself is a real trouble. The teacher should make decision about the matter of control
(objectivity of the evaluation). The holistic method cannot solve the problem of assessing unreliability if the scorer must make a judgement about the correctness of the response based on her subjective interpretation of the scoring criteria (Bachman; 1990:76). The ability to respond to an essay involves some of the writing skills such grammatical ability, lexical ability, mechanical ability (punctuation, spelling, capitalisation), stylistic skills, organisational skills judgements of the appropriacy (Kenji. K. & S. Kathleen. K.; 1999). Such skills should be taken into account when responding to students’ products. Perhaps the most difficult – and important of these skills is the judgements of the appropriacy.

The assessing procedure needs to be adapted to a logical specification of a scoring scheme. The use of descriptors for each level of such scheme can at least help make the assessment consistent and easy. One possibility is to make an analytical assessing scheme for the overall quality of the writing, but the problem is that, for example, the grammar can be good and easy to assess; which is not the case with the other writing components. It is perhaps more useful to have different sets of descriptors for each aspects of writing that you want to consider. you might want to have descriptors for grammatical correctness, use of vocabulary, content, organisation, and mechanics. These categories might be weighted differently, depending on what you want to emphasise. This is one of the most controlled way of assessing writing. It may direct the teacher’s attention to the desired criteria of evaluation; it may also raise the examiner’s doubt about his scoring procedure, but it can never solve the difficulties that trouble the testing of essays. Difficulty of assessment is still a problem even if an analytical procedure is applied.

The ease and effectiveness of assessment may be achieved if the learner’s response will be directed to the limitations set by the assignment. In other words the teacher will find no difficulties if: (1) the answer will be clearly written; (2) there will be a clear focus on the topic, and (3) if the development and focus will be achieved by means of logically and
meaningfully sequences paragraphs. Such requirements, however, appear more imaginative and unlikely to be realised. In reality, the learners’ responses can never be as the teachers’ wants, and this is the inevitable problem that he faces and which makes his assessment difficult to control.

2.4. The essay assessment approaches

To evaluate and assess their students’ essays, EFL teachers usually make call to two different procedures namely the holistic and the analytic ones. The former is agreed to be easier and quicker especially when dealing with large size classes. The latter is often described as time consuming and inappropriate in many circumstances. However, great deal of discussions and debates has been raised about each procedure’s reliability, validity and objectiveness.

2.5. The holistic approach

The assessment of an essay has different names in language testing. It is frequently referred to as subjective, holistic or impressionistic. Such a method seems to have been established independently in two similar forms in Great Britain and the United States, by Wiseman and his colleagues, and known at that time as the ‘Demon Method’ (Wiseman, 1949), and by ‘the Educational Testing’ service in the United States, better known through the work of Godshalk, Swineford, and Coffman (1966). In holistic scoring, the essays are collected from exam takers, usually responding to quite general question within a limited time. The teachers make a broad judgement of the quality of the answer in a very short time.

The holistic scoring is an approach to the whole writing assessment and not only the scoring. Its main objective is to construct what writing is, and what is important that writers should be able to do with the written language. Ideas are found to be salient trait in most contexts, but this is usually judged in a general way rather than the specific one. In other
words the teachers often check if the essay includes pertinent and convincing ideas, plenty of relevant ideas, or adequate quality of ideas, rather than the existence of the most appropriate ideas that comply with the nature of the essay’s topic. The holistic scoring, therefore, focuses only on the most prominent criteria for the context, and does not claim to assess every facet of writing competence that may appear in the students’ writing.

Such system of assessment is a subjective procedure in which the teacher makes quick judgements on writing samples and assigns an overall score. One advantage with the holistic evaluation is that it is the quickest method for scoring. The essays are read once. There is one strategy that can be used to score essay. The teacher usually uses assessing guides, which outline the features the scorer should address when scoring essays. When using assessing guides, the teacher matches features listed on the assessing guides to features on the essays and then allocates the essay the corresponding and appropriate rate. By this procedure the teacher is assumed to insure assessment reliability.

However it is not always an easy task to achieve an effective and objective assessment even with the use of the most detailed assessing guides. So, how can we then avoid or reduce unreliability in assessing? Is it possible to find a remedy to such a problematic? Is there any method which can ensure a formal and objective assessment of an essay?

2.5.1. The holistic approach deficits

There are a number of problems with holistic scoring. These problems are very serious and may cause a trouble in the field of language testing. The chief among these is that the holistic scoring is not designed to offer correction, feedback or diagnosis (Charney, 1984). The teacher’s assessment of the examiner’s work is subjective in the sense that its merit has to be evaluated by the examiner (Pelliner, 1970:19).

Such an assessment, as it is always claimed to be, rests upon reasons or principles, but the principles of assessment are truth claim in the absence of conclusive grounds. Experiences
have shown that the assessment made by a single teacher who uses the holistic method is very often unreliable. Marks are awarded on the basis of a teacher’s overall impression. Credit is not given for specific categories, and a learner’s performance is expressed as a single mark or grade. Furthermore, the students or the examinees are unable to know exactly why and how such marks are given to them. It is also possible that the teacher cannot justify the mark that he may assign to any exam-paper. If the assessment is refractory, the net result will be unreliability in scoring and this is the main factor which differentiates subjective scoring from objective one.

2.5.2. The holistic assessment new forms

No one can dare to deny that assessing an essay holistically may be unjust to students especially when it comes to tests or exams upon which depend their study success and continuity. Moreover, and even worse, during teaching sessions, holistic assessment procedure would never reflect how well students’ are progressing. Thus, the assessment essence and objectives are not attained. The teacher would never recognise his learners’ deficits and troubles with the language. Therefore, he would never have such opportunity to provide them with appropriate feedback to cure and repair the lacunae. The learners on their part would never been exposed to their mistakes, so they would stay indifferent and unresponsive vis-à-vis such impediments and hindrances. They would never have the chance to auto-correction and auto-evaluation.

Consequently researchers have attempted to enhance the situation by suggesting some methods to reduce subjectivity and increase effectiveness of the holistic assessment approach.

2.5.2.1. The administration of short essays

The single holistic scoring procedure carries backwash effects on the ground of assessment. It is therefore, impossible to achieve assessment reliability if our scoring procedures are based on an ad hoc basis. The possible and unique solution according to some
examiners is to get rid of such a type of test (that is, the single long essay questions) because they are subjective. All the methods that can be applied are presumed to be inadequate to solve the problem of unreliability in assessment. The possibility of achieving consistent and valid scoring by any rating scheme, analytic or otherwise is still a problem.

The only way to come to firm and reliable results is to avoid such a type of test. According to (Underhill; 1990: 88), one solution preferred by language test writers is to avoid subjective tests together. Such a solution has been also advocated by (Raatz, 1981) who admitted that both the oral interviews and the compositions are not tests simply because they are not objective.

As a matter of fact, some specialists propose to substitute long-type essay questions for the short essays. According to (Pelliner; 1970: 28), it is better to set short essays instead of a single long one. Such a method requires the learners to respond to more than two compulsory short essay-questions. Some teachers increase the number of questions in their exam to four short essay-questions. The time frame allocated to such an assessment is the same that is required for the single long essay-question. However, the assessment of the short essays is not the same as that of the single long essay. If the teacher allocates four short essays to his students; the general mark will be divided into four sub-marks. Five points will be allocated to each essay. The assessing procedure focuses on two criteria: language accuracy and the content (substance of writing). It is based on three steps. First, the teacher starts by reading the whole essay (that is, each essay) in order to gain a general impression of its content. The answer is either correct or incorrect. No additional information is required. The mark that should be given is from 0 to 5. The second step is referred to as an error-count method. The mark allocated to the essay can be kept as if it is free of error or it can be subtracted if it is covered by errors. The teacher may subtract, for instance, one mark from the general mark given to each essay, for each three errors. In doing so, he may differentiate between students’
abilities. The third and final step requires the teacher to count the marks given to the four essays.

However, the administration of short essays instead of a single long one has brought to light some embarrassing side effects. Evidence has shown that the inclusion of short essays as a written testing device may cause two serious problems or drawbacks.

To start with, setting up the essay questions is the most significant trouble with administrating short essays as assessment tools. Subjective assessment tests have been described as being easy in terms of construction and administration. One long question, as a fact, is not difficult to construct. However, the teacher, as an assessment designer, may find some serious nuisances in constructing more than two essay questions. It is possible to pay more attention to the construction of one question, but it seems impossible to give the other essay questions similar weight and level of difficulty. The teacher may not be able to consider two extremely important factors: the time required for testing each essay and the degree of speediness he wishes to set up into his assessment. Let assume that a maximum of one hour and half has been scheduled for the assessment. We should, then, divide ninety minutes into the number of the essay questions we wish to set forth, (let say four essay questions). Each question is assumed to take twenty-two minutes and five seconds. The problem that may happen is that one of these questions may be answered in one hour or more. In some cases, the examinee loses the designed time in order to answer just one question. He may also be perplexed by the failure to decide by which he can start. His anxiety will increase when he feels that the time is quicker than him. The assignments that he is asked to complete will be strictly evaluated with a conclusive single score.

Moreover the assessing procedure is another shortcoming of administrating short essays question. The teacher tries to find an adequate way to solve the problem of scores unreliability. His task is mainly directed to what is written by the learners whose answers are
responses to different types of questions. The short essay questions are administered to achieve two quite possible results: reliability in assessing and ease of scoring. These two factors may be achieved if no judgement is required on the part of scores (Bachman; 1990: 76). The correctness of the learner’s responses must be agreed on by different assessors, whose assessments must be identical (that is, objective assessment).

An essay, whether it is short or long, is a verbal communication. It requires the learner to use his own language in order to complete the assignment. It seems probable that the learner will be given the same score if the assessment is done by the same teacher. It is also possible that there is no chance of that happening. In point of fact, if the assessment is done one single teacher (i.e., the classroom teacher) who uses a rating scheme, the net result will be assessment reliability. However, if the assessment is quite holistic, it is generally possible to have different scores for the same exam paper. Let us assume that short essay-questions can be assessed objectively by the same classroom teacher, who prepares in advance his assessing framework. It is possible, in such a situation, to achieve an easy assessing procedure.

The single long essay-question has been rejected because it costs more time and effort than objective tests such as cloze-tests or multiple-choice items. Short essays, on the other hand, differ from long ones just in terms of length, (that is, short essay-questions require the learners to summarize their answers). The learners are assigned to respond to short essay-questions in no more than one paragraph for each question. Instructions are given beforehand so that each learner can submit his answers to the assessment’s directions. The problem which is supposed to occur is the difficulty to assess effectively and objectively all the learners’ essays over a limited span of time.

However, if we compose all the essays of one exam paper, we will get a long essay, but with different instructions. The classroom teacher will be faced with two possible factors: the length of each answer, and the differences between the topics. Each essay contains a specific
subject-matter which may differ from the others in terms of content, organization, language accuracy, purpose and writing process. Each exam paper, therefore, combines different topics and different writing processes. The differences and the length of the answers are real obstacles that the classroom teacher may encounter and which can really increase rather than reduce the difficulty of assessing.

As a matter of fact, the solution which advocates the use of short essays instead of a long one is pedagogically unacceptable. It is not possible to substitute a problem with another problem. The single long essay-question is presumed to be subjective and difficult to assess, but evidence has shown that the inclusion of short essays has entirely proved its handicap in language writing assessment. An effective assessment must be free of any problem that may hinder the assessing procedure. The attitude towards the avoidance of the administration of single long essay question because it is subjective has no sense.

There seems little reason to exclude single long essay as an assessing device simply because it is not objective. The possibility to resolve the subjectivity in assessing an essay-type examination cannot be left out. It is possible according to Nitko, (1983), Cronbach, (1984), and Gronlund, (1984) to achieve an objective scoring and to reduce scores unreliability, not by excluding the single long essay question, but by using the appropriate method. In language assessment, subjective tests such as open-ended tests permit the use of techniques that are natural and seem outwardly very valid. Some studies have shown that there are three different types of scoring methods: the Multiple-Scoring Method, the Error-Count Method, and the Analytic Scoring Method. These methods have a common objective: scores reliability, but the ability to reach such an objective is still a problematic.

2.5.2.2. The Multiple Assessing Method

The development of Multiple Assessing Method has been motivated by the desire to find ways of assessing writing with the levels of objectivity and to avoid unreliability in the
assessment of the written production. Such an attempt provides some diagnostic information to students and to their teachers that the holistic assessment can never be achieved objectively and more reliably if it is done by one single holistic assessor teacher. The teacher’s assessment is regarded as extremely subjective and unreliable. It is probable that each exam-paper can be assessed differently in different occasions. Such unreliability in scores is due to some physical and psychological factors. Such factors are clearly put forward by Heaton sees the examiner’s work as a highly subjective one based on fallible judgements, affected by fatigue, carelessness, prejudice, etc. (Heaton; 1991: 135).

Therefore, if the assessment is based on fallible judgement, the net result is rates unreliability. Thus, the only way to ensure an objective and reliable judgement of the essay is to enlist services of some equally competent assessors. In foreign language testing the definition of assessors, scorers, readers, raters, markers, or judges reflects its use for people who correct other types of language tests; they mark the papers but they never meet the individuals who write them. In a sense, the scorer is a teacher whose task is to correct the exam-papers of other students and not his own. Such a method may encourage a better scoring procedure. However, the assessor could be that teacher who tries to check how well his students are developing their writing skills.

Multiple scoring method is one aspect of the holistic assessment the theoretical foundation upon which the multiple scoring procedure is derived. Teachers make judgements of the answers as a whole: that they are unable to separate out facets of the essay and identify them. Recently, adaptations have arisen, most notably the developments of essay scales, and or rating guides to accompany the multiple scoring sessions resulting in what is known as ‘Modified Holistic Scoring’ or ‘Focussed Holistic Scoring’, but the holistic scoring still yields only one score to express the quality of the students’ essays.
The multiple assessing method implies providing separate assessments for one paper in order to obtain the suitable rate. It has been thought that it is possible that more than one assessor can in fact increase reliability than the single score of a single assessment. According to Ingram (1970: 96), the only way to increase reliability of the marking is to have several judges, whose marks are average. She noted that the judging could be perfectly adequate provided that three examiners judge each essay separately in one occasion. The teachers are asked to read the exam paper without knowing who writes it. In its classic form, the multiple scoring method consists of two readers (scorers) scoring the same script (exam-paper), but if the ratings of two readers do not agree, the paper should be read a third time, and then to accept which ever rating is nearer to the third reader. Such a procedure is used mainly in competitive examinations, and in some scoring centres.

Before each assessing session, scorers are provided with the model essays and assign a rating based on that comparison. The model essays represent borderline cases. Each essay must, by definition; fall above or below a model. One model essay represents each dividing line. The assessors should read each essay to gain a general impression of its quality in relation to the model essay. They should first make decision of the final score to be given to each essay model. Such a decision may limit ratings in some agreed marks so that no scorer can give whatever he wants. It is essential that all scorers be thoroughly familiar with the rating criteria in order to carry on a common scoring procedure.

There are four recognisable steps: First, The exam papers are read quickly by each person in the scoring group until all essays are being read. Scorers then pair up and work together to score an essay based on a scoring guide on a 6 (excellent) to 1 (poor) scale. Initially, the pair determines if a paper’s answer is on the topic or not. Papers that do not fit the ‘1’ or ‘6’ categorization are separated from the two extremes. These opposing scorings of ‘1’ and ‘6’ indicate the worst and the best scores. Such scoring is said to be initial. The next step is
recognised as a more specific scoring. It involves sorting out the essays left out of the ‘1’ or ‘6’ scoring. The readers quickly read this pile and sort the best of this middle pile into the category of ‘4’ and the worst of the pile into the ‘2’ category. Those remaining essays in the middle are the ‘3’. The final scoring is made during the fourth step. The scorers now have five distinct piles with an assigned number from ‘1’ to ‘5’ representing worst to the best as described by the traits relevance, paragraphing, style, organization, diction, mechanics, and grammar.

Multiple assessing method may have many advantages. To start with it achieves a kind of objectivity. Objectivity here refers to the maximal increase of different judgements in to maintain a solid and adequate assessment. One single judgement may affect the assessment, but more than one can be more objective and effective. The multiple scoring method is an approach to the whole writing assessment and scoring. Such a trend seeks to avoid scores unreliability and to develop an objective assessing procedure. Evidence shows that when the scores on the multiple assessing method are combined to create a single composite score in making administrative or pedagogical decisions, that single score is highly objective and reflects really the learners’ level of proficiency. The use of composite scores can increase objectivity as follow: Assume that each essay is assessed by two teachers. The result is two scores, one matched pair. We may then obtain a single by dividing the pair into two. Because two judges are used, the score will, in fact, be more objective, because it is a combination of two different judgements. Most programs also use a third scorer in cases when the first two scorers are far apart in their judgements; the way these third scores are used varies, but their result is an adjudicated score that is theoretically closer to a true score than the first two scores alone. Multiple scoring method possesses psychometric properties that enhance the objectivity of a score which can be used making yes/no decisions such as whether or not to accept the candidate into a program of study where writing competence is required and for setting cut off
points such as the level below which a student should be placed into a remedial writing program.

Moreover, increasing information about the learners’ performance is another advantage of the multiple scoring method. A key statistical question that must be resolved when using a multiple scoring method is whether scores should be combined and if so, how. If diagnostic information is part of the purpose of assessment, clearly each of the scores should be reported separately. If objectivity is a key, multiple scores when combined result in a highly objective scores. The multiple scoring method shows remarkable information of the different scores to be given to one single essay. Scores exist not simply to assign decisions, but also to communicate decisions (Hamp-Lyons, 1992). Scores are information which can be shared with the learners, their teachers, and other concerned parties and used by them to take various kinds of action in the context of the information. In contrast with the single holistic scoring where the scorer who notices an unevenness of quality in the writing has no way to report this observation, and must somehow reconcile it as a single score, multiple scoring permits judgements and differences of writing be assessed and reported.

On the other hand, the multiple assessing method may have also many drawbacks. First, the classroom assessment is always undertaken by one teacher who is in charge of the course session, the test construction and the assessing procedure. This threefold task is common for all teachers. The courses are in some cases agreed on by the administration or the curriculum and syllabuses. The testing and assessing procedures, on the other hand, is a teacher-control. Each teacher seems authoritarian in his testing and assessing. Such an authority is limited to his choice of the question-type and his method in scoring. Accordingly, the possibility to obtain help of other colleagues working on the same ground remains very low in all cases. It seems unlikely to happen that one teacher can accept to re-correct the exam-papers of other teacher. It is a truism to say that each teacher is allowed to correct just his students’ scripts.
He knows which answer is acceptable and which is not, and how exactly to deal with them. In addition, the correction may take him a long time and a great effort, because of the great number of scripts that he is confronted to. It is true that the correction of essays is very difficult and boring. Such a difficulty can be avoided if it is an objective test, such as multiple-choice items, where the answer is either right or wrong. In an essay, each teacher may regard the correctness of the learner’s answer from a slightly personal way. This may cause a disagreement among scorers over the marks they may give.

Besides, disagreement among assessors or scorers is another drawback of such multiple assessing method. There are three main causes of disagreements among scorers who independently mark the same set of scripts: first, the marks may differ in average standard or level. One marker may be generally severe, another completely lenient; Second, the marks may differ in their scatter or spread. One marker may employ the whole range of available scale, another only part of it; third, the marks may order the learners differently. According to Pelliner (1970:27), discrepancies among the markers in rank order are reflected in low intercorrelation among the arrays of marks they assign. The disagreement has a long historical background. Evidence has shown that one exam-paper can be given different scores if it is scored by different scorers. Such a ‘number paradox’, as it is referred to by Underhill (1990) can never achieve reliability on the ground that the improvement does not represent greater agreement on the value of the essay. People are inconsistent; they do not always agree, either with each other or with what they said or thought last time Underhill (1990:89). The number of markers that are supposed to avoid subjectivity and to reach a reliable and objective scoring, can in fact, increase marks unreliability. This’ hiccough’ (inconsistency) is the result of the markers’ personal judgements of the essay content, style, organization and procedure. Each marker may approach the quality of the essay from a highly subjective way. He may
regard the essay from a different point of view. He may rely on his own knowledge of the subject-matter, and how it can be responded to.

2.5.2.3. The Error-Count Method

The Error-Count Method is not new. It has had different names such as the ‘Mechanical Accuracy’ or ‘Traditional Scoring’. Earlier in this century this method was used for the purpose of achieving reliability in scoring. It has been introduced as an objective method for the assessment of an essay. The teacher’s objectives in administering such a type of test are, first; to test the students’ mastery of the courses taught, and second; to see whether or not they are able to write accurate sentences within an essay-type question paying more attention to the language form.

The procedure in scoring an essay normally goes through two steps. In the first step, the teacher proceeds by reading the whole essay in order to evaluate the quality of the answer. A general mark is, therefore, allocated to the suggested answer. The second step, on the other hand, recalls for the subtraction of marks of, hence ‘error-hunting’. The procedure, in such a situation, consists of counting the errors made by each learner and deducing the number from a given total. For example, a student may lose up to 10 marks for grammatical errors, 5 marks for misuse of words, 2 marks for punctuation; etc. The examiner in deducting the number of errors from a given total makes a distinction between two conventional types of errors: major errors and minor ones. The major or global errors are those errors which involve the overall structure of a sentence and result in misunderstanding or even failure to understand the message which is being conveyed. The minor errors, on the other hand, are those errors which cause only minor trouble and confusion in a particular clause or sentence without hindering the reader’s comprehension. They are mainly recognized in the misuse of articles, omission of prepositions, and lack of agreement between subject and verb or incorrect position of adverb. Such a distinction seems crucial in the sense it makes balance between errors. The deduction
normally follows the value of errors. According to Underhill (1990: 102), Normally, one mark is deducted for each definite error from a starting point of, for example, ten, but sometimes a distinction is made between major error (1 mark off) and minor error (½ mark off).

The Error-Count Method has been adopted to determine the effectiveness of the writing skill. The learner limits his/her focus only on two components: The relevance (the ideas of expression) and language form. The first component is used just as ‘bait’ by which the teacher can, on one hand, evaluate the student’s achievement of the courses taught, and on the other hand, to detect the types of errors that learner may make. The second component is given more attention. The accuracy of writing seems to take a major part, more than the relevance. The learner is not to be penalized for his misinterpretation or his failure to answer the question, but he will be judged for each error he will make even if he is ‘on’ or ‘off’ the topic. Such a judgement gives no interest to the purpose of the exam question. The learner’s attention is more influenced by his/her deliberate hate of errors. Everything that goes beyond such a purpose is to be considered secondary.

However the so-called Error-Count Method is not without weaknesses. Lack of validity is identified as a first shortcoming. No test can be valid without being adequate and perfect. It is only when we have a formal assessment data to look we can usefully go on to ask questions about validity. To judge validity we need to look at other kind of data. Anastasi (1982) sees construct validity is the overarching validity, and it is this type of validity which is central in writing assessment. When the test accurately measures the behaviour which defines the construct, it has construct validity. To assess content and construct validity test developers must pay more attention to the evidence for what is valued in writing in the context to which the writing test applies, design prompts to elicit that kind of writing and scoring procedures to judge those values and ensures that scores keep values in mind. These judgements of prompts
and scoring procedures are in large part content validity judgements. **Cronback (1984: 48)** calls this “logical validity”. This must be compelled with a clear sense of what is involved in the construction of written discourse, of the limitations imposed by the assessment-medium keeping in mind what it means to write in these circumstances.

The Error-Count Method has been developed to satisfy just one objective. The focus on language accuracy is, therefore, the main reason of the assessment. The other aspects of the essay writing are not taken into consideration. The Error-Count Method has proved its failure to achieve a formal assessment of an essay. It has ignored the real purpose of essay writing (communication). Construct validity, in such a situation, is completely dismissed. According to **Davies (1968)**, the validity of the test should reflect the principles of a valid theory of foreign language testing.

In addition, the essay as commonly conceived is a band of frequencies used for sending out a particular message. These frequencies are mainly organized in a way that let the message goes across. In language assessment, such frequencies are recognized as the organization of ideas, style, grammar, mechanics, handwriting, paragraphing, and so on. The message is the content, the substance of writing. If the focus is just on one of them, the assessment will lose its legitimacy. It will be considered as informal and invalid. The Error-Count Method lacks validity in two main points: the focus on language accuracy and the negative treatment of errors. The learners are assigned to respond to an essay question. Such a type of task, as a fact, needs more emphasis on the whole components of the essay rather than language accuracy.

Moreover, the learners are supposed to write an essay in a way that makes the writing process covered with definite or indefinite errors. They are asked to respond to free-stimulus and to use their own language in order to complete the task. In such a situation, the assessment will be valid if it deals with the whole rather than with the parts. The Error-Count
Method is based on the assessment on language accuracy such as grammar, vocabulary and mechanics. Such an approach is still used today by some teachers who favour to concentrate on the negative aspects of the writing task placing the learner in a position that he/she cannot write due to the fear of making mistakes. According to Ur. P., this over-emphasis on language errors can distract both learner’s and teacher’s attention from the equally important aspects of content and organization (1996:171). The essay should not be used to assess only specific components such as the mastery of language. Such an assessment is formally inadequate and informal. The scoring procedure should not be just quantitative (i.e. counting errors). Quantitative scoring procedure is inevitably impractical and some form of qualitative scoring must be formed. In language use the whole is bigger than the parts. No matter whether the parts are isolated in terms of structure, lexis or function, it is implausible to derive hard data about actual language performance from test of control of these parts alone. According to Clark, the whole of communication event was considerably greater than the sum of its linguistic elements (1983:432).

Therefore, the essay assessment should be designed to reveal not simply the language competence, but to reveal the quality of the candidate’s language performance. It is not safe to assume that a given score on the former (competence) necessarily allows conclusions to be drawn about the latter (performance). Frequently, the student’s performance and success in accomplishing the task may be masked by errors and a tired marker fail to make the necessary effort to respond to the writing as a means of communication. The assessment, according to Harris (1993:121) needs to engage first and foremost with the communicative purpose and overall coherence and organization of the text, not only with localized errors which should be a secondary concern and always related to other primary matters.

The assessor or the scorer should be more concerned with the ‘positive scoring’ Doff, (1995) which gives more emphasis to the content and meaning the learner is trying to express.
He should judge the learners for what they are able to do rather punish them for what they cannot. Writing, according to Lyons and Heasley, (1987), is presented as a problem-solving task which challenges rather than defeats the students. It is that aspect of communication that should be given more importance. The students or the examinees should not be placed in a situation where they are asked to pay more attention to language competence rather than to language performance. They should be motivated to use their own words, style and writing strategies. The scoring procedure, on the other hand, should be adapted to the communicative language testing. Obviously, this will have a better effect on the learners’ attitudes to learning.

2.6. The Analytical Assessing Method

Since most of the holistic methods, as mentioned previously, have been recognized as invalid and informal, researchers have advocated the use of a new brand of methods featured by its analytical aspect.

A single teacher is supposed to be three persons in one: classroom teacher, test designer and test assessor. He is supposed to prepare the course and to present it to the whole class; he constructs the question according to what have been already taught, and he evaluates and assesses the quality of the test answer with a conclusive mark, according to his knowledge of the subject-matter and in a way that permits the students to know exactly why and how he/she has being given such a mark. In such a situation, there is no need to enlist services of other teachers who may affect the testing procedure, and since most teachers have little opportunity to enlist services of two or three colleagues in marking class composition, the analytic method is recommended for such a purpose Heaton (1975: 137). This method depends on a careful specification of an assessing scheme, which has been carefully drawn by the teacher. Thus, one way of making subjective, impressionistic judgements more objective is to devise a marking system through bands and scales in which the judging criteria is described as precisely as possible. These bands should be made as simple as possible (range of
vocabulary, grammar, style, appropriateness, etc.) so that the teacher will not have to take into account too many aspects at the same time.

Studies such as those of Bachman (1990) strongly suggest that the assessments made by a single scorer who uses a framework (i.e., an analytic scoring system) of this sort are more reliable than the ‘global impression assessments of one person. According to him, the test such as the essay that involves the use of rating scales are necessarily objectively scored, since there is no feasible way to ‘objectify’ the scoring procedure. Bachman (1990: 76)

In fact, evidence has shown that when the scoring procedure is done by only one scorer and when the standard (i.e., the analytic scoring device) remains reasonably consistent from paper to paper, the net result is scores reliability.

2.6.1. The nature of the Analytic Assessing Method

The Analytic assessing Method comes as a reaction against the impressionistic (Holistic / Subjective) assessment. It is a psychometric method (Underhill, 1990) which is used to improve the reliability of the scoring of an essay-type examination. It depends on the Atomistic Approach (Lado, 1962), which is the breaking down of the complexities of language into isolated segments. According to Morrow (1979: 145), this influenced both what is to be tested and how this testing should carried out”.

The researchers, who advocate the use of such a method, support the arguments of the psychometric view that it is possible to reach the scoring procedure, although it is subjective in nature, can be objectively achieved.

2.6.2. The analytic assessing procedure

The procedure consists of separating the whole writing process into categories and to mark each category separately. The separate marks are, then, combined to give an overall mark to whole essay-answer. Such a ‘counting procedure’ (Pollit A, 1990) permits the scorer to limit his assessment to the ‘marking protocols’ (Underhill, 1990) that he has already selected and
graded before the administration of the test. By this way, each student is able to see how his particular grade has been obtained (Heaton, 1975: 137).

This is, indeed, a crucial point which differentiates the Analytic Scoring Procedure with the Holistic one. The teacher can easily convince his students who may ask for another correction. It is also probable that, which such a procedure, another teacher will give the same mark to the same exam-paper if he is equipped with the same assessing plan and with the same procedure.

2.6.3. The Analytic Assessing Method deficiency

The Analytic Scoring Method is not free of problems. There are, however, some frequent exertions that may not affect the scoring but the scorer himself. Studies such as those of Ingram (1970), Madsen (1983), Peter. W. Foltz, Darrel. Laham & Thomas. K. Landauer (2000) protest against the Analytic Scoring Method. They claim that the assessing procedure made by a single scorer who uses an analytic framework is entirely ad hoc and difficult to score. They, therefore, noted two existing inconveniences: (1) a problem with the specification of the assessing system, and (2) a problem with the scoring time frame.

As far as the specification of the scoring system is concerned, some teachers find that such an analytic device is quite adhoc (unplanned). It seems to them that the impressionistic approach is better and formal in use. According to Ingram (1970:96), there is no evidence that this is any more valid and reliable than the overall impression marking of experiences examiners. Such a claim is supported by Madsen (1983 ) who strongly criticizes the Wiseman method of achieving reliability on the ground that the analytic approaches to the scoring of an essay are not well-specified and do not represent greater agreements of how to weight each area of the essay. According to him, a major problem with the analytic approaches is that one never knows exactly how to weight each error, or even each area being penalized. Madsen (1983: 21).
The Analytic Method has been claimed to satisfy two interrelated approaches. The first approach is mainly recognized as The Points-Off Approach. It is based on two steps. The scorer starts first with a grade and then reduces it. In the second step, the scorer gives points for acceptable work. If we combine the two approaches we obtain the following step-by-step procedure: First, the scorer divides the general mark (20/20) into the number of components he wants to assess. Second, the scorer reads the whole essay-answer in order to gain a general impression of its content. The answer is considered either right or wrong according to the teacher’s previous knowledge of the subject-matter. Some essay-answers will be accepted because they are on the subject, and some others are to be rejected because the respondents are off-topic. Third, the scorer moves to the second step when the answer in on the subject. He tries then to evaluate the quality of the other components that he has already selected. Each of them is allotted a specific mark. Fourth, the scorer counts errors made in language accuracy (i.e., grammar, mechanics, vocabulary...) .Fifth, the scorer deduces the number of errors made by the examinee from the specific mark allotted to each area (component). Sixth, the scorer counts all the marks that he gives to all the selected components, in order to give a general mark to the whole answer.

Some teachers regard such a procedure as being invalid and time consuming. The scorer will spend more time in scoring. Such a factor, according to them, may be avoided in using the Impressionistic (Holistic) Method which is less time consuming. The scorer will be able to correct as much exam-papers as possible.

2.6.4. The analytic Assessing Method advantages

The Analytic Scoring Method, even if it is claimed to be difficult to apply, it is the most useful method that can achieve a very formal assessment. The pessimistic view raised against it has no sense in language assessment. According to Bachman and Palmer (1996), the
usefulness of a classroom assessment can be determined by considering the test’s reliability, security, and feedback. These components form the main basis of a good assessment.

The Analytic Scoring Method has been advocated in order to achieve objectivity in scoring an essay and effectiveness in assessing it. Experiences, for such a fact, have shown that one way to improve the reliability of the Impressionistic (holistic/subjective) scoring is to adopt an atomistic/analytic procedure. Such a method brings to light some considerable results that can never be realized with the other mechanical and holistic methods. The findings resulting from the application of the Analytic procedure indicate that it is possible to achieve the following advantages:

2.6.4.1. Scores reliability

Scores change from one person to another when there is no standard procedure to rely on. Reliability can be increased if we make use of an analytical assessing scheme. Such a device can firmly maintain the results. According to Pelliner (1970: 28), studies such as those of CAST (1939) strongly suggest that the assessments made by a single marker who uses a framework of this sort are more reliable than that global impression assessment made by one person”. It is possible, accordingly, to have more than one assessor giving the same score to the same exam-paper in different occasions. Such a conclusive score is due to the fact that the test scorer takes an over-restrictive view of what it is that he is testing. The scoring procedure, in such a situation yields data which is easily quantifiable.

2.6.4.2. Security

The only way to make the examinees feel at ease is to make them know how they are going to be assessed. The specification of a consistent assessing plan can easily make the classroom teacher control his marks. At the same time, each student is able to see how his particular grade has been obtained Heaton (1975: 137). Moreover, learners would take risk to write and
write again till achieving an ideal performance since they would be aware not only about their lacunae but also about the way to overcome such lacunae.

2.6.4.3. Feedback

The students know pretty how they have obtained such marks. By this way they can know in advance exactly how they have been scored and how they will be scored. The scoring system gives them meaningful feedback on various aspects of their performance. They will be encouraged to prepare language for the test and to pay more attention to how to respond to the question. The experience tells us that the students view the test as meaningful, useful, and fair if they feel that the scoring which accompanies it encourages to use their own words rather than punish them for something they do not know.

The feedback seems to be the most vital and significant outcome of an objective and reliable assessment that is why the researcher wishes to shed a slight light on such an item.

2.7. Feedback: An overview

The provision of feedback in EFL teaching and learning has been a crucial subject of controversies among language teaching theoreticians since it improves language knowledge and acquisition. It may have an enormous power of impact on learning and achievement (Hattie.2007). Before going any deeper, let’s first of all, explore the issue of definitions so that everything could be clear enough for the understanding of the terminology of such research.

In defining the term feedback, John Hattie (2008) confesses that he has struggled to understand the concept (2008.173). Furthermore, Wiggins (1997) adds that feedback is a word we use unthinkingly and inaccurately. We smile at student, say "good job" and call it feedback. We write B at the top of the paper and consider it feedback. But feedback is something different. (As cited in Burke and Pieterick.2010: 26). In fact, much of the related literature did not even attempt to define this term. Since feedback is fundamental to the
culture of learning, we need to look closer at what is meant by feedback? The terms such as impression, advice, praise, blame, and evaluation have given rise to confusion in terms of their definitions and use. However, none of these terms is feedback, firmly speaking. Therefore what is feedback? Feedback is not an advice because without feedback, no information is given on what to do to improve performance (advice) i.e. feedback is the base of an advice. Feedback is neither praise nor blame because it comprises of what worked as well as what did not work. Feedback does not mean to focus just on lacunae. Feedback is not an evaluative; it is descriptive since it is not only based on value judgments (grades, marks) about one’s performance. Finally, feedback is not simply an impression but rather a call for an action. However, Wiggins (1997) asserts that;

Feedback is something different. It is useful information about performance, it is not evaluation. Feedback is value neutral help on worthy tasks, it describes what the learner did and did not do in relation to her goals. It is actionable information, and it empowers the students to make intelligent adjustments when she applies it to her next attempt to perform.

(As cited in Burke and Pieterick, 2010: 26).

Generally speaking, feedback is the information about specific task, with a specific objective. According to Kluge & Demist (1996), feedback is seen as actions taken by external agent to provide information regarding someone’s task performance (Kluge & Demist; 1996. 235). i.e., it is an information provided by an agent being a teacher, an educational trainer, a book, a peer or any other. Hattie & Timperly clarify that it is the case that feedback is not only given by teachers, students, peers and so on, but, can also be sought by students, peers and so on, and detected by a learner without it being intentionally sought (Hattie & Timperly, 2007, 82)

Feedback is not always positive, it can also be negative. According to Dictionary.com web site, feedback is the process of returning part of output of a circuit, a system or a device to the input, either to oppose the input (negative feedback) or to aid the input (positive feedback).
Leeman sees that feedback is the information regarding both success and lacunae. (Leeman 2007:112). But, Man (2008) tackles feedback from different angles,

Feedback started in the early twentieth century with the advent of micro phones. Inputs into the mics were fed and they were designed to only work with inputs. If there were feeds that came back through the system, usually from being too close to speakers, you’d get an awful noise, that awful noise was named feedback because it was a "feed" that came back into the system.

Man (2008:153)

Man’s definition describes feedback as that awful noise usually annoying students. It is negatively seen by learners and as opposing to the input i.e. the learning process, as it may reduce and affect the students’ motivation and hinder future learning. Race also (2005) acknowledges that feedback refers to what was not achieved in a specific work. Then, feedback is given when students understand the objective and start working; they receive feedback that scaffolds their understanding and facilitate the cycle of learning.

Consequently, feedback could be defined as the all types of information provided to students regarding their performance in assessment works. It sketches the students ‘overall progress in the relevant unit. It can be provided in writing, verbally, or electronically in a number of forms to individuals, small groups, or a whole class. Therefore, feedback is intended to achieve the following objectives:

- To inform learners about their performance and achievement in assessed works and on their overall progress, either quantitatively or qualitatively.
- To draw attentions to deficits or weaknesses on their works.
- To motivate and encourage students to improve their performance.
- To assist students in their overall progress and guide them to achieve their best.

(Sid Nair & Metorva. 2013, 176)
Feedback is an essential part of formative assessment process which provides information to teachers and learners about the performance, relative to learning objectives. Yet, even much of the related literature points to the importance of feedback to student as a part of learning process. According to Black & Williams (1998), a meta analysis of assessment feedback showed that feedback resulted in positive benefits on learning and achievement across all content areas, knowledge and skill types and levels of education. In the same respect, Juwah et al. (2004:4) argue that feedback enhances knowledge and acquisition as well as it influences learning and achievement. Yorke (2002) agrees that effective feedback during the first year university can assist students’ transition to higher education. Moreover, the practice of feedback can increase students’ motivation, confidence, and enthusiasm for learning. This is because a good performance may give the students a sense of achievement which motivates them to learn more. Similarly, the identification of their proper lacunae may spur the students to multiply efforts in order to get rid of such hindrances and redress everything. Race proposes the ripples on a pond model to highlight the importance of feedback in learning.

![Race's ripple on a pond model](image)

**Figure 2.1.** Race’s ripple on a pond modal  
Adapted from Race, P. (2001)
Race identifies feedback as an essential component for successful and effective learning in the same way as many other components mainly wanting and needing, doing, and digesting. They interact with each others to achieve the following aims:

- To make sense of the task being done.
- To identify the need to learn (learning outcomes).
- To foster students want and desire to learn by increasing their confidence, enthusiasm, and self-esteem.
- To motivate learners for moving learning forward.

To conclude with, feedback is an important component in assessment, learning, and education in general. That is why, it is imperative to know about feedback and learning how to use it effectively to foster and facilitate learning and education.

2.8. The feedback effectiveness and ineffectiveness

Achieving an effective feedback is the focus or the central point of most teachers and educationalists’ investigations. Language teaching theoreticians have always sought a better quality through the effectiveness of the practice of feedback. Studies concerning this issue have examined the efficacy of the teacher’s feedback in different ways, either through investigating learners’ preferences and attitudes towards teacher’s feedback, or through the comparison of different methods and techniques used in the provision of feedback. However, feedback is still an "enigma" as acknowledged by Hattie (2012), since the findings of such studies have been inclusive and indecisive.

For the past three decades, there has been a debate concerning the effectiveness as well as the ineffectiveness of teacher feedback in the process of both teaching and learning English as foreign language contexts. Many researchers start to accept that feedback is among the most features of successful teaching and learning... and its effect are among the most variable (Hattie, 2012). Sadler (1989) claims that in order to achieve effectiveness in the provision of
feedback, several features should be considered. He approved that feedback requires knowledge of the standard or goal, skill in making much criterion comparison and the development of ways and means for reducing the discrepancies between what is produced and what is aimed for. (As cited in Taras, 2005:471). He introduced the concept of the “gap” i.e. the difference between where the learner is and where he is meant to be. In other words this gap identifies on one hand, the learners’ actual knowledge and competences along with their lacunae and hindrances, and on the other hand the knowledge and competences the learners’ are supposed to acquire at the end of a given learning unit. Thus, to achieve the efficacy in the use of feedback, teachers should have a complete understanding of "where the students are, and where they meant to be, and the more transparent they make this status for the students, the more students can help to get themselves from the points at which they are to the success points, and thus, to enjoy the fruits of feedback (Hattie, 2012). To reduce this gap between where the learner is and where he or she is meant to be. Sadler proposed two ways: Affective methods such as: Increasing learners’ motivation, efforts, and engagement. Cognitive methods: it involves reasoning, understanding...etc.

Hattie & Timperley (2007) asserts that efficient feedback can influence learning, success, and progress in general. According to Hattie & Timperley (2007), the provision of feedback is an integral part within the learning and teaching contexts. Moreover, they added that feedback is well provided when it points to mistakes related to a specific task, since they are considered as a part of the learning process. Basically, it has been indicated that an effective feedback is an essential part of the dialogue between the teacher and the learner. As Black & Williams suggest three essential principals of feedback:

- The recognition of goals: the learning outcomes should be clear to learners.
- An understanding of the learners’ current learning state.
• An understanding of a way or a method to close the gaps or discrepancies between learners’ current learning state and the desired goals.

However, providing feedback is not a free-trouble mission. Many questions about the efficiency of feedback in learning have been raised. For example, it has been stated that even a well designed feedback do not sometimes achieve its aim when accompanied with the marks. Students are most of the time interested in the quantity rather than the quality of their learning. Also, feedback may be too complex or contradictory that is misunderstood by students and hinders any opportunity for dialogue and interaction between the teacher and the student. Furthermore, the use of the feedback has the potential to be broad, and thus inappropriate and useless in overcoming the learning deficits.

In conclusion, teachers can intervene in this issue by aligning feedback with specific goals so that feedback will be more focused and directed. Williams (2011) agrees that teachers should be like coaches, they do not only focus on the current state but, they have to foster their future learning by splitting the learning into small parts so as to reach the final goals. Thus, it is useless to provide feedback of a learner current state without providing them with parameters to move the learning forward.

2.9. The feedback levels

It has been claimed that the feedback operates at four different levels:

- Feedback about the task (FT).
- Feedback about the processing of the task (FP).
- Feedback about self regulation (FR).
- Feedback about the self (FS).

The four feedback levels differ in the degree of the effectiveness they may achieve. They would work better when combined with answers to the following three questions as suggested by (Hattie & Timperly, 2007, 88):
• Where am I going?
• How am I going?
• Where to next

(The three questions will be reexamined by the researcher in the fourth chapter.)

To start with, feedback about the task deals with adding more, different, or correct information for a certain learning task or a product. In linguistic writing for instance, the teacher may use the following comment: ‘you had to highlight the type of the relationship between language and social class’. The second level is about the processing of the learning task. It is concerned with the processing of information or learning processes so as to form a product or accomplish a task. Hattie & Timperley state the following example: This page will make sense if you use the strategies we talked about earlier. Third, self-regulation feedback is concerned with self-evaluation skills to become more engaged in the task. This level can have an effect on learners’ self efficacy. In writing for example, the teacher may use the following comment: ‘you already know the features of transition from one paragraph to another, check if you have used them in your essay’. Fourth, the final level is feedback about the self. It is a personal feedback which is applied to learners as persons. Learners’ affect is the focus of such feedback level. Examples of this level are ‘good work’, ‘you are great student’ and ‘well done’. Hattie and Timperley have claimed that this level is the least effective since it is unrelated to the performance in a specific task.

2.9.1. Feedback about the task

It is the most common type of feedback. It is usually called corrective feedback. At the task level, information about, how well a task is being done, is provided. This level is related to the correctness i.e. distinguishing between correct and incorrect answers. Also, adding more different information related to a specific task. In fact, Airasian (1997) as cited in Hattie and Timperley (2007) asserts that 90% of classroom questions are directed to this information
level. However, an amalgamation of corrective feedback with information at the self level is sometimes occurred. It has been argued that FT is more effective when it is about misunderstanding, not lack of information. Another challenge is that this level cannot be generalized to other questions or other tasks. The aim of feedback is to shift students from the task level to the processing of the task, and then from processing to the self-regulation level.

However, a detailed feedback within a single level may reduce its effectiveness since it stresses not only the given objective, but also the targeted objective. Hattie & Timperley argue that:

Too much feedback only at the task level may encourage students to focus on the immediate goal and not the strategies to attain the goal. it can lead to more trial and error strategies and less cognitive efforts to develop informal hypotheses about the relationship between the instructions, the feedback, and the intended learning. (Hattie & Timperley 2007, 88)

To sum up, feedback relating to the task is more effective when it helps learners to eliminate errors and provides strategies to relate the task, the feedback, and the intended learning goals.

2.9.2. Feedback about the process

The second level focuses on the processing of the task or other tasks. It is concerned with errors detection. Thus, the role of learners here is to provide feedback for themselves. Hattie & Timperley (2007) acknowledge that such errors may indicate failure and a need to strategize, to choose different strategies to be more effective in applying strategies and or to seek help. Self-assessment is emphasized here in which learners simply strategize or asking for help, however, it depends on learners' motivation to reduce the gap between their current status and the intended one. Earley (1990) as cited in Hattie & Timperley (2007) claims that using process feedback with goal setting appears to be a direct and powerful way of shaping
an individual’s task strategy, and using outcome feedback is much less efficient way of shaping strategy. At the process level, feedback is more effective than at the task level, however, they can have an interactive effect in which the FT helps in enhancing self-efficacy and FP can bring more efficient through providing different information and finding out strategies.

2.9.3. Feedback about self-regulation

The third level is concerned with self-regulation. It involves “interplay between commitment, control, and confidence. It addresses the way students monitor, direct, and regulate actions towards the learning goals. It implies autonomy, self-control, self-direction and self-discipline (Hattie & Timperley: 2007.93). This level includes six major factors as identified by Hattie and Timperley.

These include the capability to create internal feedback and to self-assess. The willingness to invest effort into seeking and dealing with feedback information, the degree of confidence or certainty in the correctness of the response, the attributions about success or failure and the level of proficiency at seeking help (Hattie & Timperley, 2007. 94).,

Let’s have a clear introduction of each major factor or component. First, effective feedback needs to involve learners to create internal feedback and develop cognitive routines which they are engaged in through academic tasks.

For all self-regulated activities, feedback is an inherent catalyst. As learners monitor their engagement with tasks, internal feedback is generated by the monitoring process that feedback describes the nature of outcomes and the qualities of the cognitive processes that led to those states. We hypothesize that more effective learners develop idiosyncratic cognitive routines for creating internal feedback while they are with academic tasks.

However, less effective learners often include feedback to foster future learning and self-regulation strategies since they totally rely on external factors such as the teacher and the task.

Second, at this level, self-assessment is a crucial aspect in which learners should attain for an effective learning. It comprises of two aspects:

- Self-appraisal: learners evaluate their knowledge and capacities through a variety of self-monitoring processes.
- Self-management: it involves learners monitor their performance through correcting their own mistakes.

Moreover, developing meta-cognitive skills is central for learners since it assess not only their understanding, but also it extends to the strategies used in a task, and the others views about their own performance (Hattie & Timperley, 2007). They also add that as students become more experienced at self-assessment multiple dimensions of performance can be assessed. Most important, students know how and when to seek and receive feedback from others” (Hattie & Timperley, 2007:94).

Third, students’ willingness to invest efforts in seeking and dealing with feedback information is beneficial in reducing the gap between the current and the intended performance. These transaction costs include effort costs (The effort necessary for feedback search), face costs (the evaluative effects of others on individual for seeking for feedback), inference costs (the implication of inferential errors resulting from inaccurately interpreting feedback (Ashford & Cummings, 1993; De Luque & Sommer, 2000).

Fourth, the degree of confidence that learners have in the correctness of responses can affect the effectiveness of feedback. In fact, it has been noted that learners ignore feedback when the response turns to be correct, in this regard Hattie and Timperley pointed out that feedback has its greatest effect when a learner expects a response to be corrected and it turns out to be wrong (Butler, D.L., 2009:95).
Fifth, learners' attributions about success and failure can influence learners’ performance. Success-related feedback can lead to positive outcomes as well as good performance. However, failure-related feedback may result in negative outcomes, uncertain self-image and lead to poor achievements.

Sixth, seeking help can have impact on the effectiveness of feedback. It comprises of two kinds: instrumental help seeking and executive help seeking. The former is just asking for hints rather than a direct help or direct answers, and the latter is asking for direct answers to gain time and efforts. Ashford & Cummings argue that high level of proficiency of instrumental help seeking lead to self-regulation feedback. However, executive help seeking is concerned with feedback at the task or the processing of task levels. The problem with this factor that many learners do not seek help due to many reasons namely, low self-esteem and social embarrassment. (Ashford & Cummings, 1993: 89)

2.9.4. Feedback about the self

It is also called personal feedback or ego-related feedback. It is argued that this level is the least effective and the most used in classroom practices. It does not answer the feedback questions mentioned above since it contains little about the task, the processing of the task, and self-regulation activities. “Good work” or “well done” as feedback has little or no impact on achievement and learning outcomes.

To conclude, the four levels and the three questions are essential for an effective feedback and learning. It has been noted that they are related in terms of four aspects namely,

- the timing of feedback
- the effects of positive and negative feedback
- the optimal classroom use of feedback
- the role of assessment in feedback
2.10. Writing feedback

The early investigations about feedback to L2 writing focused on the feedback in language instruction, that is delivered in language class settings, for the purposes of second language acquisition (SLA), rather than on feedback to the writing skills, that is delivered in writing contexts for the long-term development of L2 writing provided in language class settings. As Ferris (2010) notices that, although L2 writing and SLA researchers look at similar phenomena, often (but not always) in similar ways, it is important to understand that they do not necessarily ask the same questions (Ferris, 2010:188).

Up until 1990s, researchers started to shed light on the role, the importance, and the effectiveness of feedback in the teaching of writing, rather than in SLA. Ferris further pointed out:

The two lines of research are not in competition, rather, they are complementary. There may be methodological gap, but it is not a philosophical or theoretical chasm. L2 writing researcher and SLA researcher who investigate written CF although they pose somewhat different questions can and should learn from each other and build on one’ another work. (Ferris, 2010:192)

It is obvious that according to Ferris, feedback is of a great impact either on second language acquisition as a whole or second language writing acquisition, since, the feedback provided to each could complete the other to achieve a better effective learning.

Interestingly, several studies have been conducted to investigate different issues concerning feedback to foreign language (FL) writing such as:

- The different ways and strategies the writing teachers could use to respond to students’ errors
- The students’ perceptions about the teacher written feedback
- And most importantly, the feedback effects on students’ writing
Drawing upon the previous and recent research studies that have been conducted, writing feedback is viewed as an essential tool for the development of foreign language writing. **Bitchener and Ferris (2012)**, who assert that feedback is an integral part in teaching writing, identify three different arguments in favor of the indispensability of practicing feedback in any writing class.

First, they have found that students, who receive writing corrective feedback (CF), revise and edit their work successfully. Here, success is defined as statistically significant reduction in the number of errors from one draft to the next (**Bitchener & Ferris, 2012: 84**). In other words, receiving feedback to writing assists students to improve their revisions and self-editing strategies for a better written production. **Ferris (2004)** sees that it can be argued that the cognitive investment of editing one's text after receiving error feedback is likely a necessary step on the road to longer term improvement in accuracy (**Bitchener & Ferris, 2012:86**).

Second, feedback to writing has an effect on student writing over time. In fact, several studies such as those of Ferris, 1995; Chandler, 2003; Ferris, 2006; Foin & Lange, 2007; and Ferris et al, 2010 have been conducted to find out the long-term effects of feedback to writing. For example, in the study of Ferris (2006), involving fifty-five ESL students in six sections of a developmental writing course at a US university, in which she compared five errors categories made in two papers: paper 1 (at the beginning of the semester), paper 4 (at the end of the semester). After a deeper review of both papers, she has reported that students who received feedback to writing made some progress in accuracy over time, but they did not correct other error categories. However, the works concerning long-term effect of feedback were criticized because of the absence of control (uncorrected) groups in the research designs, and this is a methodological problem that is hard to solve (**Bitchener and Ferris, 2012: 98**).
Third, written CF is highly valued by both teachers and students. Bitchener & Ferris (2012) argue that most writing instructors believe in and practice written CF and are unlikely to stop doing it regardless of the pronouncements of either composition theorists and applied linguists conducting experiments research (Ferris & Bitchener, 2012: 90).

Large number of studies have been conducted like those of Ashwell, 2000; Chandler, 2003; Ferris, 2006; Ferris et al, 2010 to address different issues such as the timing of written corrective feedback, the amount of feedback to writing, the form to deliver such feedback, the types of errors to correct and so on. For instance, raising the question of the best strategy to provide students with writing feedback namely the direct or indirect feedback, Chandler (2003) & Ferris (2006) find that the direct feedback led to more accurate revisions, however, Chandler reported that direct corrections results more accuracy on the next paper. Besides, Ferris has found that indirect feedback has a long-term effect on students writing accuracy (Bitchener & Ferris, 2012). Teachers should be conscious about the impacts of the feedback to writing that can help their learners to improve their accuracy and efficiency in writing, since students often trust and use the feedback provided by their teachers as previously mentioned. Several studies have been done to investigate three main issues:

- Students’ perceptions about what teacher feedback should focus on.
- Students’ preferences about the written CF
- Students preferences about form versus content feedback (Bitchener and Ferris, 2012)

Indeed, to address the question of the students’ perceptions about the focus of the teacher feedback, a large number of studies have been conducted such as Ferris, 1995; Montgomery & Baker, 2007. In the early studies, students perceive that their teachers’ feedback focus totally on the form of their writing such as grammar and mechanics, whereas, in the recent studies, the focus of the teachers’ feedback has shifted towards all the aspects of writing including, the content, organization, mechanics, and vocabulary.
Furthermore, several other studies have been conducted to raise the question of student’s preferences about writing feedback in term of which error to correct. It has been found that learners prefer writing feedback on both the content and the form, but in some they insist upon language focused feedback and in fact would be bitterly frustrated by its absence (Bitchener and Ferris, 2012).

Other studies have been undertaken to identify students’ preferences about the strategy used in the delivery of writing feedback. The researchers have compared direct and indirect feedback. They have reported that students have a slight preference of the direct correction over the indirect one. However, they do believe that the indirect corrections can be effective for long-term progress. Bitchener & Ferris claim that students prefer direct feedback on the errors because it requires less efforts for them to deal with it, but they believe that, in the long run, indirect feedback especially if errors are located and labeled with some kind of explanation or at least error codes will help them most in improving their writing (Bitchener & Ferris, 2012:94).

To sum up, there is no evidence in the studies mentioned above that the practice of feedback may have some drawbacks. Consequently, the huge numbers of research studies on feedback have all together demonstrated the vital and imperative role of feedback in both teaching and learning of writing.

2.11. The writing feedback forms

It is undeniable that feedback may exercise an immense and powerful impact upon the process of learning the writing skills. However, when it comes to the form under which such feedback should be provided, a great deal of suggestions has been set forward. Consequently, the teacher is to select the appropriate strategy according to the circumstances he finds himself in. Those circumstances often relate to the learners’ needs, their level of proficiency, their physical and psychological states of being, the learning objectives, the availability of
pedagogical aids and so. Five basic strategies for providing writing feedback can be identified whatever is the approach or the method the teacher uses.

2.11.1. Direct corrective feedback

It is the most common and used by teachers. Direct correction has been defined as a correction that not only calls attention to the error but also provides a specific solution to the problem (Bitchner & Ferris, 2012:148). It is the provision of the correct linguistic form or structure of the target language through identifying the place of an error as well as providing a specific solution to the problem. However, previous literature has investigated what is meant by "a specific solution to the problem"; it has offered three main solutions namely:

- The provision of the correct form or structure
- Putting a cross on the given error
- The addition of omitted items

This type of writing feedback is also called "Editing" as suggested by Hendrickson (1980), as cited in Bitchener & Ferris (2012). As mentioned above, direct feedback can be offered in different forms such as, crossing out incorrect or unnecessary items, rewriting, or inserting missing items and providing the correct form above the error. An example of the direct CF that has been delivered using all the three forms:

Example one

| Once upon a time, there were an old king, living with his 3 sons in the castle. The two oldest were clever. But, the 3 ones were simple and quiet. Once the king became older, he was wondering about his successor. One morning, he got a bright idea to make a challenge. The 3 brothers were supposed to bring a beautiful carpet. The most | was an three a sons
| the most |
| third was became started |
| the most |

The direct CF is useful to provide learners with explicit correction, especially when learners are beginners. They are often unable to assess themselves, or they do not know the
correct form. **Sheen (2007)** points out that this type of feedback can be effective in enhancing acquisition of grammatical structures. However, this type of feedback has its burdens. In fact, the direct correction can be frustrating for teachers as well as uncomforting for learners. **Hendrickson (1980)** argue that

One thing is certain, providing all the correct forms in students’ imperfect sentences is a time consuming ordeal that can also be frustrating to teachers, especially when they see that identical types of errors appear repeatedly on compositions written over a period of time by the same student, certainly, from the learners’ perspectives, it is disconcerting to receive a "corrected" composition with many words crossed out, new words added, and an array of marginal comments all usually written in blood red ink. Small wonder, indeed, that some students are embarrassed and lose confidence when they receive their written work corrected in this way.  

( Bitchener & Ferris, 2012:148)

Providing such kind of feedback to writing is according to Bitchener & Ferris time consuming trial. Moreover it could be so annoying and tiring for the teacher mainly when dealing with the same category of errors again and again. As for the learner, direct corrective feedback may be a source of confusion and perplexity. The learners are likely to lose confidence not only in themselves but even their teachers.

### 2.11.2. Indirect corrective feedback

Indicating the errors without providing the correction is another feedback strategy used by instructors. Indirect feedback is defined by Ferris (2002) as:

Indicating an error through circling, underlining, highlighting, or otherwise marking it at its location in a text, with or without a verbal rule reminder or an error code, and asking students to make correction themselves.  

( **Ferris, 2002: 63** )
Indirect corrective feedback localise the error but, without providing a solution. Instead, learners will correct the errors by themselves. As identified by Ferris, indirect feedback can be delivered in various ways such as: underlining, circling, highlighting, and coding. The role of learners here is to correct the errors by themselves. The same previous example will be corrected through indirect corrective feedback by underlining and using crosses.

**Example two**

<table>
<thead>
<tr>
<th>X</th>
<th>X: missing word.</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>X: wrong word</td>
</tr>
</tbody>
</table>

Once upon a time, there were old king, living with his sons in huge castle. The two oldest were clever, but the ones were simple and quiet. Once the king older, he wondering about his successor. One morning, he got bright idea to make challenge. The 3 brothers were supposed to bring beautiful carpet.

It has been noted that indirect feedback is effective to long-term learning of the target language since it challenge the learner to correct their errors by themselves and reflect upon them. Pollard (1990) argues that the role of the teacher using the indirect CF is a merely a "reflective agent", who can provide their learners with hints about the place and the type of an error. In this regard, learners can relate these hints to the context, identify the place of or the type of an error, and correct it based on their background knowledge. Besides, language research shows that this type of feedback is more beneficial than the direct CF since it engages learners in the correction process and the use of different cognitive skills to move their learning forward. Many studies has reinforced this claim, for instance, Chandler’s (2003) study, which involved thirty-one second language learners at university, shows that indirect correction through underlining learner's errors is preferable to direct correction. Another study by Ferris (2006), which involved ninety two ESL learners in the USA, providing them with both direct and indirect correction, indicates that the teacher’s indirect feedback help learners to revise their drafts and editing them successfully.
2.11.3. Metalinguistic corrective feedback

It is a feedback strategy that involves providing learners with explicit corrective comments. In this regard, teachers should provide clear information about the place of an error, the nature of an error and how to correct it. Ellis et al (2006) add that explicit corrective comments can be provided in two different forms.

The first and the most common is the use of error codes by addressing different kinds of errors with abbreviated labels. The same previous example will be used to show how corrective feedback is provided through error codes.

Example three

Once upon a time, there were old king, living with his sons in the castle. The two oldest

\[
\begin{array}{lll}
\text{s-v agr} & \text{m.art} & \text{w.w} \\
\times \text{were clever. But, the ones were simple and quiet. Once the king } & \text{become older, he was} \\
\text{m.n} & \text{w.w} & \text{w.ch} \\
\text{wondering about his successor. One morning, he got bright idea to make challenge. The three brothers were supposed to bring the beautiful carpet.} \\
\text{m.sup}
\end{array}
\]

In the above example, some errors abbreviations are suggested. Therefore, ‘s.v agr’ , ‘m.art’, ‘w.w’, ‘w.t’, ‘m.n’, ‘w.s’, ‘w.ch’, and ‘m.sup’ stand respectively for subject-verb agreement, missing article, wrong word, wrong tense, missing noun, wrong structure, wrong word choice, and missed superlative.

Various studies have been conducted to compare the provision of error codes with other types of written corrective feedback. One of them was conducted by Lalande (1982), he finds that a group of second language learners, who were offered error codes correction, improved their writing’s accuracy. However, the other group that was provided direct correction commits more errors (As cited in Ellis.2008). Another study conducted by Ferris (2006) reveals that error codes correction does assist learners to improve the accuracy of their writing.
The second form of metalinguistic CF is to identify the place of an error with some sort of explanations about the given grammatical or linguistic rule. The same example will be corrected using the grammatical explanations.

**Example four**

| Once upon a time, there **were** × old king, living with his 3 sons in the castle. The two oldest × were clever. But, the 3 ones were simple and quiet. Once the king **become** older, he was wondering about his successor. One morning, he got × bright idea to make challenge. The three brothers were supposed to bring the × beautiful carpet. |
| (1) Subject-verb agreement. The use of the past simple with “to be” -‘was’ is used with singular pronouns, and ‘were’ with plural pronouns. |
| (2), (4), (9) (5) The use of articles. There should be an “a” before the noun when it is mentioned for the first time. However, there should be a "the" before the noun when such a noun has been mentioned previously. |
| (8) You need to use the past simple for narrating past events. |
| (3) & (6): Numbers should be written in letters. |
| (9) Inappropriate word / ‘Once’ indicates the start of an action. |
| (11) Missing word. |

This type of CF is very useful since it provide learners with explicit information about the nature of the error and the grammatical or the linguistic explanations. In a study realized by *Ellis et al (2006)*, involving thirty-four ESL learners in New Zealand, two types of corrective feedback were compared (explicit feedback with metalinguistic explanations and implicit feedback) with a control group receiving no feedback at all. They have found that explicit feedback with metalinguistic explanations assist better students in enhancing their writing.
However, this type of feedback is rarely used by teachers, this is due to either that it is obviously too much time-consuming, or because it calls for the teacher to possess sufficient metalinguistic knowledge to be able to write clear and accurate explanation for a variety of errors. (Ellis, 2008:101).

2.11.4. Focused versus unfocused feedback

Much of the literature stressed the idea of whether to correct all the errors in learners’ written production such as, content, ideas organization, grammar and mechanics or just select specific errors to be corrected. In fact, focused feedback differ from the unfocused one in that the former deals with the correction of all the types of errors whereas the latter is concerned with a selective correction, that is to say, the teacher is going to select the errors to be corrected. For instance, in the same example mentioned before, the teacher could have chosen the article errors for the correction process. Ellis argues that processing corrections is likely to be more difficult in unfocused CF as the learner is required to attend to a variety of errors and thus, is unlikely to be able to reflect much on each error (2008.102). Correcting all students’ errors is likely to be time consuming task as well as frustrating for both teachers and learners.

Bitchener & Ferris (2012) argue that early studies (Kepner, 1991; Robert et al, 1986; Semke, 1984) concerning the effectiveness of CF and except of Van Beuningen (2008), have failed due to the fact that the type of the feedback used was unfocused.

Focused CF is considered more effective than the unfocused one as the learner is able to examine multiple corrections of a single error and thus, obtain the rich evidence they need to both understand why what they wrote was erroneous and to acquire the correct form (Ellis,2008:102). Several studies were conducted to investigate the effectiveness of focused corrective feedback. One of them was done by Sheen (2007); she investigated the use of English articles by ninety-one ESL learners in the USA. She used two treatment groups: direct correction only, direct metalinguistic correction groups, and a control group. She has found that those in the treatment groups who receive focused CF performed better than the control
group. As mentioned above, it has been proved that focused feedback has a greater effect in enhancing accurate language use of the grammatical features.

2.12. The teacher’s roles in delivering feedback

To respond to students writing is a challenging task for teachers as it requires them to negotiate different rhetorical and teaching purposes which in turn ask them to inhabit different response personas (Deirdre Burke, 2010). The question to be asked here is what roles do teachers perform when responding to students writing?

It has been hypothesized that writing teachers should adopt four different roles when checking students writing based on feedback focus, purpose, and the timing of feedback. (Burke & Pieterick, 2010). According to investigations, the writing teacher is to perform different roles mainly those of a reader, a writing teacher, a grammarian and a judge.

To start with, the teacher is primarily a reader before anything else. The teacher is expected to read and react to the text. However, reading students’ texts are not an easy task because teachers “must concentrate intensely and must analyze, synthesize, and interpret simultaneously” (Burke and Pieterick, 2010:31). To respond to students writing, the teacher should answer this question “How do I react to a written text as a reader?”. The role of the teacher here is to respond to the content of the text without providing any corrections or suggestions. For instance, he could say “This is a good idea, I agree with you ”. Second, another role of the teacher is a writing teacher who acts as a guide, ask questions, clarify certain ideas in students’ writing without giving any corrections. For example, “what do you mean by this?” Third, the writing teacher is also a grammarian. As its name implies, the role of the teacher here is to identify the grammatical errors concerning the use of tenses, subject verb agreement and etc... The teacher can also provide some grammatical explanation trying to treat the students’ common errors. Fourth, the writing teacher is a judge or an evaluator. This role is very important since it is essential to inform learners about their performance. In
brief, the teacher is an amalgamation different and various roles towards teaching, testing, assessing and feedback the learners’ writing abilities.

2.13. Conclusion

The search for an objective assessing procedure which can achieve a conclusive and a reliable assessment is still a problem that should be resolved. The essay, even if it is a very valid and efficient testing device, has contributed greatly to our understanding of the effects of subjective judgements on the assessing procedure. Subjective examination such as the essay has brought to light a very serious problem that impels teachers to adopt different strategies in order to avoid it. Such a problem, known in language assessment as scores unreliability, is one of the main result of the single holistic scoring or more precisely the subjective scoring of a written discourse. While there is nothing that anyone can do to avoid such a problematic, there are some methods that can be applied to reduce the effects of subjective judgements.

The inclusion of the four types of the scoring methods, such as the Essays-Questions type Method, the Multiple-Holistic Scoring Method, the Error-Count Method, and the Analytic Scoring Method, into the field of testing written language, such as the essay-type examination, have really clarified the problem of subjectivity in scoring an essay-type examination. Such methods work around a common point - avoiding or reducing subjectivity judgements in the hope to achieve a more exact and reliable assessing procedure – but they have not all reached their purpose.

The possibility to come to an objective assessment as it is the case with objective tests can be achieved just by the Analytic Method, which is quite quantified in a mathematically precise sense. Such a method is assumed to be the only formal type of methods which may maintain validity in relation with reliability. The other methods, on the other hand, have not been satisfactory enough. Some of them have lacked validity, even if they have come in some ways to conclusive results. Some others have lacked reliability, even if they have avoided
subjectivity. Objectivity in scoring an essay-type examination can be maintained if there is a high probability to bridge the gap between both reliability and validity. Such an attempt is likely to be realised if we make an adequate use of the Analytic Scoring Method, which is claimed to be as the more appropriate testing method that may reach a more conclusive and more formal assessing procedure.

An effective assessment would certainly result in an effective feedback able to enhance and redress the EFL learners’ writing performance. Feedback is crucial aspect in education. It is not only limited in the cycle of foreign language learning and teaching. But, it also extends to assist learners to enhance their writing skill. Feedback is defined as useful information related to a specific task and objective. To achieve the effectiveness of feedback, teachers and learners should ask three main questions, that provide the feed up, feedback, and feed forward, which works at four different levels. It is obvious that providing feedback on students' writing is a challenging task. The teacher should have the knowledge of different forms of feedback and different aspects such as, the timing and the amount of the feedback. It is also needed from the teacher to perform different roles to check students' writing.
Chapter Three

Scrutinizing the findings
# Chapter Three

## Scrutinizing the Findings

3.1. Introduction

3.2. The Second Year Written Test Assessment

   3.2.1. The Researcher’s Assessing Procedure
   
      3.2.1.1. The Success and Failure Rates Reliability
      3.2.1.2. The Researcher’s Assessment Unreliability

   3.2.2. Students' Assessing Procedure
   
      3.2.2.1. The Sampling Frame
      3.2.2.2. The Students’ Assessment Contribution

3.3. The Third Year Written Test Assessment

   3.3.1. The Impressionistic Assessing Procedure

   3.3.2. The Researcher’s Assessment
   
      3.3.2.1. The Unreliability of the Passing / Failing Marks: Group A
      3.3.2.2. Unreliability of the Assessing Procedure:

   3.3.3. Essays’ Samples Analysis

3.4. Teachers’ Contributions (Multiple Scorings)

3.5. The Teachers’ Interview

3.6. The Results’ Interpretation

3.7. Conclusion
3.1. Introduction

As formerly declared the chief purpose of the present investigation is to suggest some useful approaches, methods and techniques towards achieving and enhancing the teachers’ assessment to their EFL learners writing. Such an issue has been tackled several times by a lot of educators and scholars. To contribute to such a field, the researcher attempts, first to prove that the human assessment to the learners writing is not always effective and reliable, and second to suggest an alternative assessment through the computer. The researcher’s suggestion is based upon an elaboration a highly designed assessing architecture. In other words the research wants to suggest a kind of cooperation between the teacher and the computer. The teacher is to decide upon the criteria to assess in any written production by elaborating an assessing scheme and the computer is to put into practice such scheme. The computer would apply such an assessing scheme superhuman precise way.

Therefore, the researcher uses considerable experimental tools in order to maintain or to check the credibility of the propose hypothesis. The analysis of the different assessments of the two populations’ written tests in addition to the results of the questionnaire point out to the existence of different factors affecting the teachers’ writing assessment. The inclusion of the written tests as experimental tools may provide an answer to the unreliability of the teachers’ assessments to writing. Moreover it could determine which kind of assessment the EFL learners’ need to enhance their writing skills. The obtained results are used as experimental teaching for the purpose to set forth a very efficient assessing method.

The analysis of the data is undertaken in three different steps. In the first step, the assessments of the secondary population written tests have been fully analyzed. The contribution of both teachers and students have reinforced the credibility of the analysis in a way that permits a very useful assessment of the students’ writing. In the second step, the findings of the tests’ assessments are compared with the findings of the questionnaire. The
final step, however, is devoted to a sharp analysis of the main population written tests assessments. Chapter three is so linked to the first one, since it deals with the population studied and the tools used. The researcher goes back to the research questions and the hypotheses he put forward about the ineffectiveness and unreliability of the teachers’ assessments to writing.

3.2. The second year written test assessment

Testing writing usually means to evaluate the different components of the writing skills such as the relevance of content, the diction, the style, the organisation of ideas, the coherence and cohesion of the written piece, the conventions of writing, and so on. However, the test administrated to the second year EFL students as part of their first semester examination was merely meant to check and evaluate how well they have seized and understood the nature of the human language its features and functions. Such a test was considered as a guided writing which is usually reliable and objective to score.

On the other hand, the second year students were considered as a control group. They were given a cloze test as a part of their exam process on Tuesday, January 28th, 2014. Such a test format (see appendix 1) had been designed in order to evaluate students’ abilities to decide on the correct items. The test was mainly used to prove that objectivity and reliability are not exhausting aims to achieve when it comes to written cloze tests.

3.2.1. The researcher’s assessing procedure

The classroom teacher is the researcher himself. He corrected 60 exam papers of the two groups in less than three days. Such an assessment had been undertaken in three continually occasions. The following table displays the marks obtained by the second year students after the researcher’s first assessment procedure. The students in each group are classified in two different tables according to the marks they got, i.e., those students with passing marks ($\geq 10$) in a single table, and those with failing marks ($< 10$) in another.
<table>
<thead>
<tr>
<th>Students</th>
<th>The first assessment</th>
<th>The second assessment</th>
<th>The third assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>2.</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>5.</td>
<td>14,5</td>
<td>14,5</td>
<td>14,5</td>
</tr>
<tr>
<td>6.</td>
<td>14,5</td>
<td>14,5</td>
<td>14,5</td>
</tr>
<tr>
<td>8.</td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>10.</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>11.</td>
<td>13</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>13.</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>9.</td>
<td>13,5</td>
<td>13,5</td>
<td>13,5</td>
</tr>
<tr>
<td>10.</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>20.</td>
<td>13,5</td>
<td>13,5</td>
<td>13,5</td>
</tr>
<tr>
<td>22.</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>23.</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>25.</td>
<td>12,5</td>
<td>12,5</td>
<td>12,5</td>
</tr>
<tr>
<td>26.</td>
<td>11</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>28.</td>
<td>11,5</td>
<td>11,5</td>
<td>11,5</td>
</tr>
<tr>
<td>30.</td>
<td>11</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>31.</td>
<td>12,25</td>
<td>12,25</td>
<td>12,25</td>
</tr>
</tbody>
</table>

Table 3.1. Second year written test results (Passing marks /Group A)

<table>
<thead>
<tr>
<th>Students</th>
<th>The first assessment</th>
<th>The second assessment</th>
<th>The third assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.</td>
<td>03</td>
<td>03,5</td>
<td>03,5</td>
</tr>
<tr>
<td>4.</td>
<td>03</td>
<td>03</td>
<td>03</td>
</tr>
<tr>
<td>7.</td>
<td>09,5</td>
<td>09,5</td>
<td>09,5</td>
</tr>
<tr>
<td>9.</td>
<td>05</td>
<td>05</td>
<td>05</td>
</tr>
<tr>
<td>12.</td>
<td>08</td>
<td>08</td>
<td>08</td>
</tr>
<tr>
<td>16.</td>
<td>07</td>
<td>07</td>
<td>07</td>
</tr>
<tr>
<td>17.</td>
<td>06</td>
<td>06</td>
<td>06</td>
</tr>
<tr>
<td>18.</td>
<td>08,5</td>
<td>08,5</td>
<td>08,5</td>
</tr>
<tr>
<td>19.</td>
<td>04</td>
<td>04</td>
<td>04</td>
</tr>
<tr>
<td>21.</td>
<td>0,25</td>
<td>01</td>
<td>01</td>
</tr>
<tr>
<td>24.</td>
<td>06,5</td>
<td>06,5</td>
<td>06,5</td>
</tr>
<tr>
<td>27.</td>
<td>04</td>
<td>04</td>
<td>04</td>
</tr>
<tr>
<td>29.</td>
<td>04,25</td>
<td>04,25</td>
<td>04,50</td>
</tr>
<tr>
<td>32.</td>
<td>09,5</td>
<td>09,5</td>
<td>09,5</td>
</tr>
</tbody>
</table>

Table 3.2. Second year written test results (Failing marks /Group A)

It is obvious from the ‘Table 3.1.’ that the assessment procedure is so consistent and reliable since the teacher has provided almost the same passing marks in the three different assessment occasions. However, the ‘Table 3.2.’ has revealed a kind of unreliability and
inconsistency in the group A failing marks allotted by the researcher mainly to the students in red, numbers: 03, 21, and 29. Most surprisingly the same thing happened with the researcher when assessing group ‘B’ i.e. the passing marks were so consistence to achieve reliability whereas some failing marks were not so. The table below exhibits the marks obtained by the second year students after the researcher first assessment.

<table>
<thead>
<tr>
<th>Students</th>
<th>The first assessment</th>
<th>The second assessment</th>
<th>The third assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>10.5</td>
<td>10.5</td>
<td>10.5</td>
</tr>
<tr>
<td>6.</td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>7.</td>
<td>12.5</td>
<td>12.5</td>
<td>12.5</td>
</tr>
<tr>
<td>10.</td>
<td>11.5</td>
<td>11.5</td>
<td>11.5</td>
</tr>
<tr>
<td>11.</td>
<td>11</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>13.</td>
<td>14</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>15.</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>16.</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>17.</td>
<td>13.5</td>
<td>13.5</td>
<td>13.5</td>
</tr>
<tr>
<td>18.</td>
<td>11</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>19.</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>22.</td>
<td>10.5</td>
<td>10.5</td>
<td>10.5</td>
</tr>
<tr>
<td>23.</td>
<td>13.5</td>
<td>13.5</td>
<td>13.5</td>
</tr>
<tr>
<td>24.</td>
<td>13.5</td>
<td>13.5</td>
<td>13.5</td>
</tr>
<tr>
<td>26.</td>
<td>11</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>27.</td>
<td>11</td>
<td>11</td>
<td>11</td>
</tr>
</tbody>
</table>

Table 3.3. Second year written test results (Passing marks / Group B)

<table>
<thead>
<tr>
<th>Students</th>
<th>The first assessment</th>
<th>The second assessment</th>
<th>The third assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.</td>
<td>08</td>
<td>08</td>
<td>08</td>
</tr>
<tr>
<td>3.</td>
<td>0.25</td>
<td>01.5</td>
<td>01.5</td>
</tr>
<tr>
<td>4.</td>
<td>01</td>
<td>02</td>
<td>02</td>
</tr>
<tr>
<td>5.</td>
<td>07</td>
<td>09</td>
<td>09</td>
</tr>
<tr>
<td>8.</td>
<td>02</td>
<td>04</td>
<td>04</td>
</tr>
<tr>
<td>9.</td>
<td>7.5</td>
<td>07.5</td>
<td>07.5</td>
</tr>
<tr>
<td>12.</td>
<td>04</td>
<td>04</td>
<td>04</td>
</tr>
<tr>
<td>14.</td>
<td>05</td>
<td>05</td>
<td>05</td>
</tr>
<tr>
<td>20.</td>
<td>05</td>
<td>05</td>
<td>05</td>
</tr>
<tr>
<td>21.</td>
<td>04.5</td>
<td>05</td>
<td>05</td>
</tr>
<tr>
<td>25.</td>
<td>06</td>
<td>06</td>
<td>06</td>
</tr>
<tr>
<td>28.</td>
<td>01.5</td>
<td>01.5</td>
<td>01.5</td>
</tr>
</tbody>
</table>

Table 3.4. Second year written test results (Failing marks / Group B)
It is so visible from the ‘Table 3.3.’ above that the marks were consistent and reliable in a mathematically precise way, unlike ‘Table 3.4.’ that displays many inconsistent allocated failing marks for the same written piece. The fact points to easiness and simplicity of assigning a passing mark and therefore identifying a successful exam. However when it comes to an average written exam, the assessor generally could not make his mind. Allocating a mark or assessing such rank of written paper would be so complicated. It is always so difficult to assess and evaluate an average written test whether it was a cloze test or an essay-type test. On the other hand, assessing a well written test would be effortless. The trouble would remain only on what mark could better suit such a well written answer.

All in all, the whole results obtained throughout the three different assessing procedures of each group demonstrate two possible facts: first, the reliability of the passing and failing rates and second, the unreliability of the assessing procedures.

### 3.2.1.1. The success and failure rates reliability

In group ‘A’, the passing marks (i.e. 10 and more) of the three different assessing procedures are 18. That is to say, in each assessing occasion there are 18 students who are assumed to pass the exam. However, there are 14 students who are assumed to fail the exam in each assessing occasion. They failed the exam because they obtained less than 10 in their exam as it is clearly shown in the following table.

<table>
<thead>
<tr>
<th>Scoring procedures</th>
<th>Passing marks</th>
<th>Failing marks</th>
<th>Total points</th>
</tr>
</thead>
<tbody>
<tr>
<td>First assessment</td>
<td>18</td>
<td>14</td>
<td>288.5</td>
</tr>
<tr>
<td>Second assessment</td>
<td>18</td>
<td>14</td>
<td>289.75</td>
</tr>
<tr>
<td>Third assessment</td>
<td>18</td>
<td>14</td>
<td>290</td>
</tr>
</tbody>
</table>

Table 3.5. Group “A” success rates

The possibility to reach assessment reliability is very high. Both passing and failing scores seem inconsistent. The number of the students who are assumed to pass or fail the exam is kept the same from scoring occasion to another. The validity of such results is clearly shown.
through the following graph. The graph displays the rate of success and failure in each assessing procedure of group A.

Graph 3.1. The group (A) success and failure rates
The ‘Graph 3.1.’ shows clearly that passing and failing scores seem consistent from an assessing procedure to another. In each assessing occasion the same number is allotted. But this does not mean that the teacher’s assessment is completely reliable.

3.2.1.2. The researcher’s assessment unreliability
On the other hand and if we count up all the points that were given to 32 written subjects in each assessment; we can deduce a slight difference between the assessing procedures. Such a difference is clearly shown in the graph below which displays the total points allocated to group (A) students members in each assessing attempt.
According to the above graph, there is a slight difference between the three different procedures. There was a low degree of assessment unreliability. Such unreliability can be recognised in the following scores exhibited in table below. The assessing procedure was affected by the unreliability of three scores. The table below shows that three scores out of thirty-two had been changed thus causing the unreliability of assessment.

<table>
<thead>
<tr>
<th>Students' numbers</th>
<th>The first assessment</th>
<th>The second assessment</th>
<th>The third assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>03</td>
<td>03,5</td>
<td>03,5</td>
</tr>
<tr>
<td>21</td>
<td>0,25</td>
<td>01</td>
<td>01</td>
</tr>
<tr>
<td>29</td>
<td>04,25</td>
<td>04,25</td>
<td>04,5</td>
</tr>
</tbody>
</table>

Table 3.6 The assessment unreliability

According to the above table the students’ marks had changed during the second and third time of the assessment procedure. The written productions of the students 3 and 21 had received extra points in the second assessment. The same marks had been kept in the third assessing procedure. However, the written essay of the student 29 had kept the same mark in the second assessing procedure but it had had a slight change in the third assessment. Such unreliability can be shown in the following graph which displays the unreliable marks in group (A) in each assessing attempt.

![Graph 3.3. The marks unreliability in group (A)](image-url)
The numbers mentioned in the graph are out of 100. That is to say, 300 is equal to the score 03 and so on. The graph needs big numbers in order to demonstrate clearly each value. According to this graph, the assessing procedures show clear but not extravagant unreliability.

- The exam paper ‘3’ was allocated the score 03 (300 in the graph) in the first assessment process. But such a mark had become 03, 50 (350 in the graph) during the each of the two following assessment occasions. The difference between the two allocated marks is 0, 5. It seems that the second and third assessment procedures were more reliable, consistent and objective than the first one.

- The exam paper ‘21’ was allocated the score 0, 25 (0,250 in the graph) in the first assessment process. However, the mark had shifted to 01 (100 in the graph) in the following assessment occasions. Consequently, the second and third assessment procedures seemed more reliable and consistent than the first one.

- As opposed to the two previous written tests, the third exam paper had undergone a clear change in the third procedure. The mark allocated to the student ‘29’ remained the same in the first and the second assessment procedures (4, 25). But, during the third attempt it had become 4, 50 with a 0, 25 extra point. Therefore, the reliability, that seemed to take place in the assessment of the two previous paper (i.e. papers ‘3’ and ‘21’) during the first and second procedures, was lost when assessing the exam paper ‘29’.

Counting the whole points allocated to the whole population in the first and third assessments, 01, 5 extra points mark the difference between the two procedures. Such a difference can also be calculated in such way: The total point of the third scoring minus the total points of the first scoring. It can be found as follow: $290 - 288.5 = 01.5$

Similarly, in group B, the passing and failing rates were kept the same from one assessing occasion to another. There are (16) students who are assumed to pass the exam in each
assessing occasion as opposed to (12) students who are down marked. The difference is clearly seen when we combine all the points given in each assessing occasion. The unreliability of the assessing procedures can be maintained as follow: The total points given in the third procedure minus the total points given in the first one. Consequently, following result could be obtained: $244 - 237.25 = 6.75$

Comparing the results obtained in group A to those obtained in group B, the researcher can summarise that in group A the total points changed from assessing procedure to another as it is indicated in ‘Table 3.3.’; However, in group B such a change in mainly recognised in the second attempt. The results of the third procedure are kept the same as in the second one. In the first scoring the total points is 237.25. In the second and the third assessing occasion, the total point is 244. The difference is 06.75 points.

The difference between the results of group A and those of group B can recognised in the number of the written productions that had witnessed a slight change. In group A the unreliability of the assessing procedure can be seen in three written essays; however in group B there are five written essays which are assumed to have caused unreliability.

All in all, the probability to reach assessment reliability is very high. The results show clearly that only 08 out of 60 written essays have witnessed slight assessment unreliability. The other marks obtained in the 52 other exam papers have maintained their consistency from an assessing procedure to another. The fact is mainly due to ease of assessing such kind of objective test. The answer is either right or wrong. However, such kind of tests tells a little about the students’ achievements of the learning objectives. In other words, neither an effective feedback nor a corrective session could be designed from the part of the instructor. The test would never achieve the summative assessment aim. The graph below points up clearly the rate of percentage between the reliability and the unreliability of the cloze test.
results through displaying the total points allocate to group (A) in each assessing attempt.

Graph 3.4. The degree of reliability

According to the previous graph, there is a reliability of 86.66 percent as opposed to unreliability which appears very low with a percentage of only 13.33 percent. The assessing procedure had been affected by extra factor such as the illegibility of the words. There had been some problems to understand some words. Such illegibility in writing is the main factor that makes the researcher assesses again and again the same written essay. The conclusive marks, however, are the ones which have been given in the third procedure. This means that there is no reason to doubt about the final marks that had been given to the students. Such marks would determine the students’ failure or success. The students were submitted to such marks when they were asked to correct their own exam answers. That was the second step that the researcher adopted in order to check the reliability of the assessment he made and to make students know exactly why they had been allotted such rates.

3.2.2. Students' assessing procedure

The students were also invited to have a contribution to such study. The contribution of the students to the assessing process was very important for the researcher since to may give another dimension to the issue under investigation. In other words, students are undeniably a very significant component of the whole process of teaching, testing and assessing, therefore
they should be involved in every step towards achieving effectiveness and success. Moreover, the researcher aimed at raising his students’ awareness about the process of writing assessment. The researcher selected one essay and three students from each group to assess those essays. Such selections were based on the following sampling frame.

3.2.2.1. The sampling frame

The sampling frame, as demonstrated in the ‘Figure 3.1.’ below, had been used in order to select one essay from each group and six students who were assumed to correct and assess the two selected written productions. The selected essays are those of the student number 25 in group A and of the student 21 in group B as indicated in the following table.

<table>
<thead>
<tr>
<th>Group A</th>
<th>Selected essays</th>
<th>1st assessment</th>
<th>2nd assessment</th>
<th>3rd assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>25</td>
<td>12,5</td>
<td>12,5</td>
<td>12,5</td>
</tr>
<tr>
<td>Group B</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>04,5</td>
<td>05</td>
<td>05</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.7. The selected essays samples

It is worth to mention that the selection of the two essays was done in purpose. The essay ‘25’ was selected to check up the consistency of its three assessment procedures. However, and unlike the previous selection, the essay ‘21’ was chosen because of the inconsistency of its three assessment processes. Three students were also randomly selected from each group to be the assessors of the selected written papers. The selected students in group A are the members whose numbers are numbers are (19), (13) and (16). Whereas, the selected students in group B are the members whose numbers are (20), (23) and (26).

The selected written essay in group ‘A’ was allotted the letter ‘A’ in order to represent the group A and the students were allotted the numbers 1, 2 and 3. In group B, the selected essay paper was allotted the letter ‘B’ in order to represent the group ‘B’, and the students assessors were allotted the numbers 4, 5 and 6 as it is shown in the table below.
Table 3.8. The selected students as assessors

The figure below recapitulates the whole procedure adopted in order to select the written paper samples and the students to contribute to the process of assessment.
3.2.2.2. The students’ assessment contribution

The students’ contributions to the assessing procedure have gone through the following three attempts: self-assessing, assessing exchange and multiple-assessment. The results of the students’ contributions are as follow:

In the self-assessing attempt, the results obtained show that both students A and B gave the same marks as those given by classroom teacher or the researcher in his third assessing occasion. The student who represents A whose number is 25 allocated 12, 5/20 to his own written production. The student B whose number is 21 granted 05/20 to his own exam paper.

In the second attempt or the assessing exchange attempt, the results obtained were the same as in the previous attempt despite the exchange of written papers. The student A gave 05/20 to the essay of the student B, and the student B gave 12, 5/20 to the paper of the student A.

Finally, the last attempt or the multiple-assessment procedure is mainly regarded as the most important one in determining the degree of reliability of the assessing process. The six selected students were invited to assess the two selected written pieces. The table below includes all the results that had been obtained through the assessment of the six students to the written papers A and B.

<table>
<thead>
<tr>
<th>The essays</th>
<th>Student 1</th>
<th>Student 2</th>
<th>Student 3</th>
<th>Student 4</th>
<th>Student 5</th>
<th>Student 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>12,5</td>
<td>12,5</td>
<td>12,5</td>
<td>12,5</td>
<td>12,5</td>
<td>12,5</td>
</tr>
<tr>
<td>B</td>
<td>05</td>
<td>05</td>
<td>05</td>
<td>05</td>
<td>05</td>
<td>05</td>
</tr>
</tbody>
</table>

Table 3.9. The students’ assessments to their peers’ writings

The students’ contributions as it is demonstrated above confirm the teacher’s assessing procedure. It is almost a truism to say that the marks granted by the eight students, the students A and B and the six others, to the essays A and B are similar to those provided by the researcher. The table below exhibits a clear agreement over the marks or the assessment of essay papers A and B provided by the whole participants.
The essays

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>The researcher final mark</td>
<td>12.50</td>
<td>05</td>
</tr>
<tr>
<td>The student A mark</td>
<td>12.50</td>
<td>05</td>
</tr>
<tr>
<td>The student B mark</td>
<td>12.50</td>
<td>05</td>
</tr>
<tr>
<td>The student 1 mark</td>
<td>12.50</td>
<td>05</td>
</tr>
<tr>
<td>The student 2 mark</td>
<td>12.50</td>
<td>05</td>
</tr>
<tr>
<td>The student 3 mark</td>
<td>12.50</td>
<td>05</td>
</tr>
<tr>
<td>The student 4 mark</td>
<td>12.50</td>
<td>05</td>
</tr>
<tr>
<td>The student 5 mark</td>
<td>12.50</td>
<td>05</td>
</tr>
<tr>
<td>The student 6 mark</td>
<td>12.50</td>
<td>05</td>
</tr>
</tbody>
</table>

Table 3.10. The whole assessment procedure

The following diagram describes better the final assessment provided by the entire participants to the written cloze-test. The participants are the researcher himself, the two students whose written papers were selected for such an assessment, and six more students selected to contribute to the procedure i.e. nine people were involved in the process of assessing the two cloze-test samples.

![Graph 3.5. The final assessment provided by the whole participants](image)

The final marks of the cloze test question

The different assessing attempts

Paper A

Paper B

Written papers A and B

1 2 3 4 5 6 7 8 9

Graph 3.5. The final assessment provided by the whole participants
This graph consists of two lines: vertical and horizontal line. The vertical line to the left hand includes the two different samples of assessment (written production A and written production B). The essay A is represented by the long colon and the exam paper B is represented by the short one. Each colon of both written papers, in turn, represents an assessing procedure. The obtained marks are clearly shown on the top of the different colons. Each mark is multiplied by ten so that the graph can be clearly established.

According to this graph the results obtained by the different assessors follow a consistent level. The students’ assessing procedures seem to meet the results that had been given by their teacher. Although they had no experience in such field of testing, they had reached the same possible results. The degree of reliability seems to have reached its top. They found no difficulties in assessing their peers’ responses. Such possibility to ensure a consistent assessing procedure in mainly recognised in objective tests. The results are precisely set forth. They are supposed to be applied for the testing of more formal sciences such as physics and mathematics.

The representative sample of close test question shows clearly that the assessing procedure is determined to satisfy the following criteria: marks reliability, ease of interpretation and ease of assessing. As it is clearly confirmed in the previous graphs, the marks remain consistent from an assessing procedure to another. The objectivity of such a test is quite possible to achieve. The assessing procedure, in such a situation, is approached to be objective because marks can be maintained even by different assessors. It is neatly perfect in the case of cloze test where the teachers seem very often like a machine. Even the students who had no experimental learning in the assessing field were able to be formally conclusive in their assessing procedures. The assessing procedure reached a very high degree of reliability.

Moreover, the marks are perfectly obtained in less than three days. The interpretation of the test responses was very easy. There had been just some problems of handwriting which did
not affect the assessing process. The problem of illegibility of some exam papers caused minor frustration in the assessing procedure. This can be even recognised in more formal sciences. The students could read their own handwriting but they can never assume to be correct if the test answers are either correct or wrong. The test answers need no additional assistance on the part of the students who are not asked to use their own words and styles in order to complete the task.

The test is objective in the sense that it permits no personal information. *Passing a test*, according to (Morrow; 1979:145) involves making a specified proportion of correct responses. The assessing procedure, therefore, is limited to whether the students’ responses are right or wrong. No judgement, which can affect marks, is required on the part of the teacher. Even the students were able to judge the quality of each other’s responses. The cloze test answer as it is demonstrated in (appendix B) is restricted to some limited items. The students were equipped with such an answer in order to see whether they are right or wrong. There are no alternatives that can be used.

Finally, in such a test, the probability to assess many papers without any trouble is very high. If the test is based on some specific words that the students are asked to complete, the assessing procedure will be automatically easy. The instructions that are given in the exam questions as it is demonstrated in appendix A can help both the classroom teacher and the students to see whether the answers are right or wrong. This means that the test is clearly set forth in advance in a way that can make each student forecast his mark is equipped with the correct answer.

### 3.3. The third year written test assessment

Third year students were considered as the main population and the experimental group too. They were given an essay-type test as part of the examination process, on Sunday, January 26th, 2014. The students are asked to discuss, i.e., to present the different aspects of
the suggested test question (see Appendix 2), in a limited time of about an hour and half. The assessment of such written test had gone through different processed namely the impressionistic or holistic assessment and the analytic assessment.

3.3.1. The impressionistic assessing procedure

Such an approach of assessment is based upon a general evaluation of the essay without taking into account any criteria. Therefore, it seems to be so subjective. The teacher’s assessment in this case is always affected by various factors such as his opinion about the subject matter, his prejudices about the leaner himself, his actual state of being during the process of assessing and so on.

3.3.2. The researcher’s assessment

The researcher spent 15 days to assess 92 exam papers. It was difficult to state very precisely which criteria were to be tested. Each exam paper was to be judged according to the teacher’s feeling and prejudice. The assessing procedure adopted was quite holistic. Credit was not given to specific criteria. The researcher encountered serious problems in his assessing procedure. The researcher assessed the whole essays in three different occasions with a rest of three days between each assessment and another. The assessing procedure was based upon the following criteria. According to the researcher, there are seven levels upon which an essay may be evaluated: excellent, good, high average, low average, weak, very weak, off topic.

First, the essay is judged as “Excellent” when the mark is about 13.50/20 to 14/20. The essay demonstrates superior writing ability. It is distinguished by its fully and effectively developed content and by its rhetorical cleverness. An essay in this category may have the following attributes:

- The essay identifies the main points of discussion and analyses each one in a thoughtful manner.
The essay supports each point with insightful reason and examples.

The essay develops its ideas in a clear, organised manner, with appropriate transitions to help connect ideas together.

The essay demonstrates proficiency, fluency, and maturity in its use of sentence structure, vocabulary, and style.

The essay demonstrates an excellent command of the elements of Standard English, including grammar, word usage, spelling, and punctuation but main contain minor flaws in these areas.

The criteria for lower scores are the same as the one suggested above; the only difference is that the standard for quality decreases for successively lower scores.

Second, the essay is evaluated as “Good” when the mark is from 12/20 to 13/20. The essay demonstrates strong writing ability. It may be less thoughtful or less polished than the first rank, but it will be solid in content and development and will employ an effective style. It is clearly a passing essay. An essay in this category has the following attributes.

- The essay goes beyond a routine response to the topic while satisfying all the terms of the assignments.
- It has a focused thesis or main idea and a clear and appropriate organization.
- It is fully developed using specific examples and details.
- It demonstrates clear language control and a general facility of diction, phrasing, and sentence structure.
- It may have minor flaws or occasional awkwardness, but it will be largely free of serious errors in mechanics, grammar, and usage.

Third, the essay is rated as “High average” when its mark is about: 11. 50 and 10. 50

The essay demonstrates adequate college-level writing ability. It may be undistinguished in content, development, or style, but the writing is competent enough to indicate that the writer
is ready for an upper-division writing course. Language weaknesses do not significantly limit
the writer’s ability to develop and communicate ideas. An essay in this category typically may
have the following attributes:

- It may respond somewhat routinely or simplistically to the topic, but it satisfies all the
terms of the assignments at least minimally or implicitly.
- It has a recognizable main idea and an apparent organization, however, mechanical
uses some specific details or examples to develop and clarify ideas.
- It demonstrates basic competence in diction, phrasing, and sentence structure,
although there may be some imprecision, clumsiness, and /or repetition.
- It has minor errors in mechanics, grammar, and usage, but these will be neither
frequent nor serious enough to confuse or significantly distract the reader.

On the other hand the following criteria constitute a no credit judgement. They are referred
by failing rates. They constitute the four remaining essay levels.

First the essay is assigned as “Low average” if the mark is about 09/20 to 10/20. The essay
reflects inadequate college-level writing ability. It is marked by significant weaknesses in
content, development, or expression, indicating that the writer is not prepared to handle upper
division writing. Language weaknesses significantly limit the writer’s ability to develop and
communicate ideas. An essay in this category typically has one or more of the following
weaknesses:

- It may be confused or inconsistent in its response to the topic, or fail to satisfy the
most important terms of the assignment.
- It is confused, unclearly or incoherently organized, or logically flawed.
- It lacks sufficient specific details and examples to clarify or develop ideas.
- It has errors in mechanics, grammar, and usage that are frequent and serious enough to
confuse and distract the reader.
Second, the written production is assessed as “Weak” when the allocated mark is about 08./20 to 06/20. This is a type of essay which represents weak college-level writing. It is marked either by severity of weakness or by the combination weaknesses in development and expression that make the writing nearly incoherent. It is clearly not a passing essay. An essay in this category typically has the following weaknesses:

- It may seriously confuse or misconstrue the assignment.
- It has obvious and significant flaws in organization and/or logic.
- It lacks specificity, either of examples or reasons, or its use of details is largely irrelevant.
- It lacks control of diction, phrasing, and sentence structure. It has such frequent and serious errors in mechanics, grammar, and usage that the writing is largely incoherent and meaning is neatly lost.

Third, the written piece is evaluated as “Very weak” when the mark is about 03. 50/20 to 05/20. The essay represents the most minimal response to the assignment. It is almost completely unsatisfactory. Such an essay is so slight that the writer appears to have been overwhelmed by the assignment, or so error-hidden that it is incoherent.

Finally, the essay is estimated “Off-topic” when its mark is about 0. 50/20 to 03/20. All the answer is wrong. The student does not understand the question. Such an essay receives from 0.50/20 to 03/20 according to the quality of language competence the writer exhibited.

This is at least what researchers have agreed upon. However, everything is to be based upon the teachers’ estimations and appreciations of the circumstances they may find themselves in. Consequently, and before going any further, it is advisable to display the marks allocated to both groups in each assessing occasion. The results of both groups will be followed by a deeper and full analysis. The following table displays the results obtained by the students of group ‘A’ after the three assessment attempts provided by the researcher himself.
<table>
<thead>
<tr>
<th>Students’ numbers</th>
<th>1st assessment :/20</th>
<th>2nd assessment :/20</th>
<th>3rd assessment :/20</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>11</td>
<td>10</td>
<td>09,5</td>
</tr>
<tr>
<td>2.</td>
<td>06,5</td>
<td>08</td>
<td>10</td>
</tr>
<tr>
<td>3.</td>
<td>12</td>
<td>12,5</td>
<td>13</td>
</tr>
<tr>
<td>4.</td>
<td>13</td>
<td>11</td>
<td>12,5</td>
</tr>
<tr>
<td>5.</td>
<td>08,5</td>
<td>09</td>
<td>10</td>
</tr>
<tr>
<td>6.</td>
<td>07</td>
<td>05</td>
<td>06</td>
</tr>
<tr>
<td>7.</td>
<td>07</td>
<td>05</td>
<td>05,5</td>
</tr>
<tr>
<td>8.</td>
<td>08,5</td>
<td>06</td>
<td>07</td>
</tr>
<tr>
<td>9.</td>
<td>04,5</td>
<td>05</td>
<td>06</td>
</tr>
<tr>
<td>10.</td>
<td>13</td>
<td>14</td>
<td>13,5</td>
</tr>
<tr>
<td>11.</td>
<td>03</td>
<td>01</td>
<td>02,5</td>
</tr>
<tr>
<td>12.</td>
<td>08</td>
<td>06</td>
<td>07</td>
</tr>
<tr>
<td>13.</td>
<td>10</td>
<td>09,5</td>
<td>10,5</td>
</tr>
<tr>
<td>14.</td>
<td>08,5</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>15.</td>
<td>07,5/20</td>
<td>06,5/20</td>
<td>08/20</td>
</tr>
<tr>
<td>16.</td>
<td>05</td>
<td>03</td>
<td>03</td>
</tr>
<tr>
<td>17.</td>
<td>13,5</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>18.</td>
<td>08</td>
<td>09</td>
<td>08,5</td>
</tr>
<tr>
<td>19.</td>
<td>06</td>
<td>05,5</td>
<td>05</td>
</tr>
<tr>
<td>20.</td>
<td>06,5</td>
<td>07</td>
<td>05</td>
</tr>
<tr>
<td>21.</td>
<td>11,5</td>
<td>12</td>
<td>12,5</td>
</tr>
<tr>
<td>22.</td>
<td>10</td>
<td>10,5</td>
<td>11</td>
</tr>
<tr>
<td>23.</td>
<td>07</td>
<td>08</td>
<td>08,5</td>
</tr>
<tr>
<td>24.</td>
<td>11,5</td>
<td>10,5</td>
<td>10</td>
</tr>
<tr>
<td>25.</td>
<td>08,5</td>
<td>09,5</td>
<td>10</td>
</tr>
<tr>
<td>26.</td>
<td>10,5</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>27.</td>
<td>03,5</td>
<td>05</td>
<td>04</td>
</tr>
<tr>
<td>28.</td>
<td>04</td>
<td>05</td>
<td>03,5</td>
</tr>
<tr>
<td>29.</td>
<td>12</td>
<td>12,5</td>
<td>11,5</td>
</tr>
<tr>
<td>30.</td>
<td>08,5</td>
<td>10</td>
<td>09,5</td>
</tr>
<tr>
<td>31.</td>
<td>13</td>
<td>14</td>
<td>13,5</td>
</tr>
<tr>
<td>32.</td>
<td>07</td>
<td>10</td>
<td>09</td>
</tr>
<tr>
<td>33.</td>
<td>05</td>
<td>08,5</td>
<td>07</td>
</tr>
<tr>
<td>34.</td>
<td>05</td>
<td>06</td>
<td>07,5</td>
</tr>
<tr>
<td>35.</td>
<td>07,5</td>
<td>10</td>
<td>09,5</td>
</tr>
<tr>
<td>36.</td>
<td>12</td>
<td>11</td>
<td>11,5</td>
</tr>
<tr>
<td>37.</td>
<td>06</td>
<td>05</td>
<td>04,5</td>
</tr>
<tr>
<td>38.</td>
<td>14</td>
<td>13,5</td>
<td>13</td>
</tr>
<tr>
<td>39.</td>
<td>06</td>
<td>08,5</td>
<td>07</td>
</tr>
<tr>
<td>40.</td>
<td>13</td>
<td>12,5</td>
<td>13,5</td>
</tr>
<tr>
<td>41.</td>
<td>02</td>
<td>03</td>
<td>02,5</td>
</tr>
<tr>
<td>42.</td>
<td>12</td>
<td>11,5</td>
<td>11</td>
</tr>
<tr>
<td>43.</td>
<td>14</td>
<td>12,5</td>
<td>13</td>
</tr>
<tr>
<td>44.</td>
<td>13</td>
<td>12</td>
<td>12,5</td>
</tr>
<tr>
<td>45.</td>
<td>03</td>
<td>01</td>
<td>02,5</td>
</tr>
<tr>
<td>46.</td>
<td>06</td>
<td>05,5</td>
<td>07</td>
</tr>
<tr>
<td>47.</td>
<td>13</td>
<td>12</td>
<td>12,5</td>
</tr>
</tbody>
</table>

Table 3.11. The group ‘A’ marks after the three holistic assessment attempts
The second following table displays the results obtained by the students of group ‘B’ after the three attempts of assessment of the classroom teacher or the researcher.

<table>
<thead>
<tr>
<th>Students’ numbers</th>
<th>1st assessment :/20</th>
<th>2nd assessment :/20</th>
<th>3rd assessment :/20</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>11</td>
<td>11.5</td>
<td>10.5</td>
</tr>
<tr>
<td>2.</td>
<td>04</td>
<td>05</td>
<td>04,5</td>
</tr>
<tr>
<td>3.</td>
<td>02,5</td>
<td>04</td>
<td>03</td>
</tr>
<tr>
<td>4.</td>
<td>09,5</td>
<td>10.5</td>
<td>10</td>
</tr>
<tr>
<td>5.</td>
<td>12</td>
<td>13</td>
<td>12,5</td>
</tr>
<tr>
<td>6.</td>
<td>06</td>
<td>08</td>
<td>07,5</td>
</tr>
<tr>
<td>7.</td>
<td>14</td>
<td>13</td>
<td>13,5</td>
</tr>
<tr>
<td>8.</td>
<td>11</td>
<td>12</td>
<td>10,5</td>
</tr>
<tr>
<td>9.</td>
<td>05</td>
<td>02</td>
<td>02</td>
</tr>
<tr>
<td>10.</td>
<td>01</td>
<td>01</td>
<td>01</td>
</tr>
<tr>
<td>11.</td>
<td>05,5</td>
<td>06</td>
<td>05</td>
</tr>
<tr>
<td>12.</td>
<td>13</td>
<td>13</td>
<td>12,5</td>
</tr>
<tr>
<td>13.</td>
<td>10</td>
<td>09,5</td>
<td>11</td>
</tr>
<tr>
<td>14.</td>
<td>04</td>
<td>07</td>
<td>06</td>
</tr>
<tr>
<td>15.</td>
<td>11</td>
<td>10,5</td>
<td>11,5</td>
</tr>
<tr>
<td>16.</td>
<td>11</td>
<td>10,5</td>
<td>09,5</td>
</tr>
<tr>
<td>17.</td>
<td>05</td>
<td>07,5</td>
<td>07</td>
</tr>
<tr>
<td>18.</td>
<td>05,5</td>
<td>08</td>
<td>07</td>
</tr>
<tr>
<td>19.</td>
<td>11</td>
<td>12</td>
<td>11,5</td>
</tr>
<tr>
<td>20.</td>
<td>09</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>21.</td>
<td>12.5</td>
<td>11,5</td>
<td>12</td>
</tr>
<tr>
<td>22.</td>
<td>09</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>23.</td>
<td>10</td>
<td>09</td>
<td>11</td>
</tr>
<tr>
<td>24.</td>
<td>11</td>
<td>10</td>
<td>09,5</td>
</tr>
<tr>
<td>25.</td>
<td>10</td>
<td>09</td>
<td>10</td>
</tr>
<tr>
<td>26.</td>
<td>09,5</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>27.</td>
<td>14</td>
<td>13,5</td>
<td>12</td>
</tr>
<tr>
<td>28.</td>
<td>07,5/20</td>
<td>08/20</td>
<td>09,5/20</td>
</tr>
<tr>
<td>29.</td>
<td>03</td>
<td>05</td>
<td>02</td>
</tr>
<tr>
<td>30.</td>
<td>00</td>
<td>05</td>
<td>01</td>
</tr>
<tr>
<td>31.</td>
<td>12,5</td>
<td>11</td>
<td>11,5</td>
</tr>
<tr>
<td>32.</td>
<td>10</td>
<td>12</td>
<td>10,5</td>
</tr>
<tr>
<td>33.</td>
<td>14</td>
<td>14,5</td>
<td>15</td>
</tr>
<tr>
<td>34.</td>
<td>02,5</td>
<td>01</td>
<td>01,5</td>
</tr>
<tr>
<td>35.</td>
<td>0,5</td>
<td>01</td>
<td>01,5</td>
</tr>
<tr>
<td>36.</td>
<td>13</td>
<td>12</td>
<td>12,5</td>
</tr>
<tr>
<td>37.</td>
<td>10</td>
<td>09,5</td>
<td>08,5</td>
</tr>
<tr>
<td>38.</td>
<td>09</td>
<td>08</td>
<td>09,5</td>
</tr>
<tr>
<td>39.</td>
<td>12</td>
<td>13</td>
<td>11,5</td>
</tr>
<tr>
<td>40.</td>
<td>10</td>
<td>08,5</td>
<td>09</td>
</tr>
<tr>
<td>41.</td>
<td>05</td>
<td>06</td>
<td>07</td>
</tr>
<tr>
<td>42.</td>
<td>07</td>
<td>06,5</td>
<td>09</td>
</tr>
<tr>
<td>43.</td>
<td>01</td>
<td>02,5</td>
<td>03</td>
</tr>
<tr>
<td>44.</td>
<td>14</td>
<td>13</td>
<td>12,5</td>
</tr>
<tr>
<td>45.</td>
<td>12</td>
<td>11,5</td>
<td>12,5</td>
</tr>
</tbody>
</table>

Table 3.12. The group ‘B’ marks after the three holistic assessment attempts
The results of the three different assessing procedures of each group show two possible significant facts: the unreliability of the passing and failing marks and the unreliability of the assessing procedures. They can be interpreted as follow:

**3.3.2.1. The unreliability of the passing / failing marks: Group A**

In group A, the passing marks (i.e. $\geq 10$) of the three different assessing procedures differ from an assessing occasion to another. In the first procedure 19 students passed the exam and 28 failed; in the second assessing setting 22 students passed the exam and 25 students failed. In the third assessment the same results of the first one had been given as shown in the following table.

<table>
<thead>
<tr>
<th>Assessing attempts</th>
<th>Passing marks</th>
<th>Failing marks</th>
<th>Total marks</th>
</tr>
</thead>
<tbody>
<tr>
<td>First assessment</td>
<td>19</td>
<td>28</td>
<td>403,5</td>
</tr>
<tr>
<td>Second assessment</td>
<td>22</td>
<td>25</td>
<td>410,5</td>
</tr>
<tr>
<td>Third assessment</td>
<td>19</td>
<td>28</td>
<td>415</td>
</tr>
</tbody>
</table>

**Table 3.13. Group A passing and failing rates**

The previous table shows that the passing and failing marks change from an assessing circumstance to another. Such a change happens in the second assessment. In the third assessing setting the same number of students who pass and fail the exam is kept as in the first scoring. The graph below shows a slight frustration of assessment through the unreliability of the researcher’s holistic assessment.

**Graph 3.6. The unreliability of the passing and failing mark**
The graph shows that the marks which represent passing and failing the exam change from the first assessment to the second. The third assessment is kept in the same level as the first. The unreliability can be clearly demonstrated if we count all the points that are given in each assessing attempt. As it is referred to in Table 3.1, there is a clear difference between the total points of the three assessing attempts. Such a difference is represented through the graph which displays the unreliability of the total points allocated by the researcher in the three different attempts.

Graph 3.7. The unreliability of the total points

Unlike the Graphs 3.5 and 3.6, the Graph 3.7 demonstrates a high degree of unreliability in terms of the total amount of points given in each assessing setting. The assessing procedure seems unreliable in the sense that the marks allotted changes their weight from down to up. In the first assessment attempt the researcher provided a sum of 403, 5 points to the whole population of group A; in the second attempt, the total had been raised of 07 points, and of 05 points in the third assessment setting. The difference can be seen as counted as follow:

\[ 415 - 403.50 = 12.50 \]

3.3.2.2. Unreliability of the assessing procedure

The assessing procedure is affected by the unreliability of all the marks allotted to 47 written papers. The marks change their weight from an assessing setting to another. The graph
below revealed the unreliability of marks given to ten written papers in the three different assessing attempts. The graph displays the unreliability of the assessing procedures in the three different attempts.

**Graph 3.8. The unreliability of the assessing procedures**

The preceding graph consists of a vertical line which refers to the three assessing attempts and a horizontal line that refers to the ten written papers. The numbers above each colon refers to the given mark. Each number is multiplied by ten in order to clearly set up the graph. It displays the mark given in each assessing try. According to the graph the mark allocated to each written production in each assessing procedure appears inconsistent. As an example, in the assessment of the number one essay, it is visibly recognisable the gap between the three different procedures as represented in the following graph. It displays the unreliability of assessment of the exam paper one as an example.
Graph 3.9. The marks’ unreliability

As it is confirmed in the previous graph, the written essay 1 had been given three different marks in different assessing instances. In the first assessment procedure the student had been allotted 11/20. In the second and in the third assessment, the first mark had been decreased of 01, 5. In some other written papers, however, the marks obtained in the first assessing step had been increased in the third assessing procedure. Such a difference as it is referred to in the diagraph can be also seen in all the marks given to each written papers. The marks are either increased or decreased.

As far as group ‘B’ is concerned, approximately the same analysis and interpretation may be said. In group B, the same marks’ unreliability had been revealed. In addition to the unreliability of the assessing procedures, the failing and passing rates show similar marks frustration. Counting the percentage of success and failure in the examination through the three assessment procedures, the same rates would be found. However, it does not mean that the assessment was so objective and reliable if we have a look to the dissimilar marks allocated by the teacher in each of the assessing occasion. The main difference between the
two groups can be summarised in the table below which displays the number of the passing and failing marks in group B.

<table>
<thead>
<tr>
<th>Assessing attempts</th>
<th>Passing scores</th>
<th>Failing scores</th>
<th>Total points</th>
</tr>
</thead>
<tbody>
<tr>
<td>First assessment</td>
<td>23</td>
<td>22</td>
<td>381</td>
</tr>
<tr>
<td>Second assessment</td>
<td>22</td>
<td>22</td>
<td>399</td>
</tr>
<tr>
<td>Third assessment</td>
<td>22</td>
<td>22</td>
<td>388</td>
</tr>
</tbody>
</table>

**Table 3.14. Group B passing and failing rates**

According to the previous table the failing marks in group B are kept standard from an assessing circumstance to another. The passing rates, however, show a slight change in the second assessing attempt, but the final number is limited to 22 students in both the second and the third assessment. The difference between the three attempts is clearly seen if we count all the points given in each assessment step. The graph displays the unreliability of the passing and failing marks in the three attempts of assessment.

**Graph 3.10. The marks’ unreliability (Group B)**

According to the graph both the second and the third assessing attempts share the same level. The difference is mainly recognised in the first procedure where the level of the passing marks is upper than the failing. The difference becomes very large if we compare the results
of the total points of the first and the second assessing procedure. Such a difference can be seen as follow:

- The difference between the first and second procedure there has reached eighteen 18 points:

\[399 - 381 = 18\]

- The gap between the second and the third procedure is about eleven 11 points:

\[399 - 388 = 11\]

- Finally, the difference between the third and the first procedure is seven 7 points:

\[388 - 381 = 7\]

The unreliability of the assessing procedure can generally recognise in the assessing of the whole written essays of both groups. The population consists of 92 students who are divided into two categories. The category A represents those who pass the exam and category B represents those who are assumed to fail the exam. The table below underlines the general differences between the three differences. Such differences are mainly recognised in the passing / failing rates and the total points that contribute greatly to our understanding of assessment unreliability. The table exhibits the unreliability of the passing and failing marks of the whole population of the study

<table>
<thead>
<tr>
<th>Assessing procedures</th>
<th>Passing marks</th>
<th>Failing marks</th>
<th>Total points</th>
</tr>
</thead>
<tbody>
<tr>
<td>First assessment</td>
<td>42</td>
<td>50</td>
<td>784.5</td>
</tr>
<tr>
<td>Second assessment</td>
<td>44</td>
<td>47</td>
<td>809.5</td>
</tr>
<tr>
<td>Third assessment</td>
<td>41</td>
<td>50</td>
<td>803</td>
</tr>
</tbody>
</table>

**Table3.15. The unreliability of the success rates of the whole population**

The information on the previous table are clearly explained through the following graph. The graph displays the unreliability of the assessing process used for the whole population. It
displays the passing marks, the failing marks, and the total allocated points in the three attempts of assessment.

![Unreliability of the assessing procedure](image)

**Graph 3.11. The assessing procedure unreliability**

The previous diagram can interpret assessment unreliability in two main approaches: First, the passing and failing marks do not remain static. They go up and down. In other words, the number of the students who are assumed to pass or fail the exam changes from an assessing procedure to another. Second, the unreliability of the number of the students who pass or fail the exam is due to the unreliability of the assessing procedure itself. Marks are not maintained standard from assessing procedure to another. The researcher increases and reduces the weight of each exam paper in each assessing attempt. The amount of points added or reduced from each exam paper increases or decreases the number of the passing / failing scores.

### 3.3.3. Essays’ samples analysis

The teacher’s assessing procedure seems inconsistent. The threefold attempt has shown clearly that marks are affected by personal opinions. The teacher’s judgement of the quality of each essay answer seems to frustrate from assessing attempt to another as in it clearly noticeable in some students’ marks in bold type in **Table 3.10**. Such factor is really
maintained by concrete facts. The teacher’s doubt over the mark to be allocated to each written paper increases from assessing attempt to another. The written productions number 15 in group A and number 28 in group B would serve as perfect samples for an effective analysis as displayed in the following table. The table points out the selected samples for a deeper analysis.

<table>
<thead>
<tr>
<th>Written papers</th>
<th>First assessment</th>
<th>Second assessment</th>
<th>Third assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>A (15)</td>
<td>07,5</td>
<td>06,5</td>
<td>08</td>
</tr>
<tr>
<td>B (28)</td>
<td>07,5</td>
<td>08</td>
<td>09,5</td>
</tr>
</tbody>
</table>

**Table 3.16. The selected essays samples**

According to Table 3.14., the marks given to both written papers change from an assessing procedure to another. Although the criteria of evaluation that the classroom teacher or the researcher relied on in his assessment and which the constituted the basis of the method adopted, the marks had not been controlled enough. The assumption is that two or more written essays had not to be given the same scores urge the classroom teacher to read again each answer. Such an assumption is due to the teacher’s doubt over the quality of the answer that he was unable to evaluate or assess with a conclusive mark. If we consider, for instance the first marks of both written essays we can see clearly that they had been given the same mark. In the second assessing attempt, however, a slight change occurs. The written paper A lost 01 point as opposed to the written paper B which had acquired extra point. The doubt over the marks given to both written papers in the second assessment continued when the exam paper A was given 08/20 in the third assessing occasion. The exam paper B, however, had been automatically gained extra point. Assessment unreliability is clearly verified in the graph below which displays the unreliable marks allocate to the written essays samples.
Graph 3.12. The essay samples marks unreliability

The graph shows a clear difference between the marks provided in the different scoring procedures. The differences could be summarised as follow.

- In the first assessing procedure both the written papers A and B receive the same level. According to the researcher’s assessing scheme such types of essays are recognised as weak. They deserve failing marks such as 07, 5/20.

- The difference is very apparent in the second assessing where a slight change occurs. The evaluation of the exam paper B had gone up as opposed to the exam paper A which was down marked. Both essays, however, are recognised as weak even if there is a slight change.

- The essay B in the second assessment seems on the same level as A in the third assessment. Both are recognised as failing rates. They are weak essays.

- The appreciation of the written paper B was raised in the third assessment attempt. Such a type of essay is recognised as low average essay because the given mark for such an essay is between 10/20 and 09, 5/20.

As it is demonstrated above the unreliability of the assessing procedure is seen as a combination of two interrelated factors: falling/raising marks and falling and rising of the
grading criteria. Such factors are the results of the unreliability of the assessing scheme that the classroom teacher adopted in the assessing processes. It is quite true that there is a zero percent of reliability. The degree of unreliability is very high; whereas, the degree of reliability appears completely low.

Since no essay paper has received the same mark in the three processes of assessment, it can be deduced that the assessing unreliability has reached the extreme rate of 0%. In a more scientifically sense: this can be explained as follow: zero (0) reliability (R) multiplied by percentage value (100) and divided by ( / ) the number of the exam papers (92) the result is equal to zero percent.  

\[ 0 \times \frac{100}{92} = 0\% \]

Or; this can be also explained as follow: the unreliability of the marks in each assessment multiplied by the percentage value and divided by the number of the exam papers the result will be equal to one hundred percent of unreliability.  

\[ 92 \times \frac{100}{92} = 100\% \]

The unreliability of the results is explained in the two graphs below which display the total amount of points in each assessment step.

**Graph 3.13. The three assessments results**

The following graph displays the total amount of points in the first assessment step usually seen as reliable.
According to Graph 3.13, there are differences in the points given to the exam papers. The amount of points changes from an assessing procedure to another. Such a change clearly refers to marks unreliability. The degree of unreliability is determined by a careful analysis of the three different assessing attempts. If we compare the results obtained in the first assessing with the second and third ones, we support the approach of assessment unreliability in an essay test. The first assessing attempt can be used as model of comparison (Graph 3.13). According to the first results, the marks obtained in the second and third assessing procedure are commonly conceived as unreliable. The first assessment is always presumed to attain 100% of reliability (Graph 3.14). In Graph 3.15 below the degree of unreliability is clearly shown.

**Graph 3.14. The first assessment results reliability**

**Graph 3.15. The degree of unreliability**
According to the previous Graph 3.15., there is an unreliability of 100%. Such a percentage can be explained through some basic mathematics.

784.5 points in the first assessing attempt represents 100% of reliability
784.5 (1st assessment) – 809.5 (2nd assessment) = (–25)
784.5 (1st assessment) – 803 (3rd assessment) = (–18.5)
\[\Rightarrow (-25) + (-18.5) = 43.5\]

\[\text{a-} (-43.5) \quad \Rightarrow 100\% \quad (-25) \quad ? \quad (-25) \times 100 / 43.5 = (-57\%) \text{ of unreliability.}\]

-That is to say the degree of reliability goes down of about 57% in the 2nd assessing attempt is concerned.

\[\text{b-} (-43.5) \quad \Rightarrow 100\% \quad (-18.5) \quad ? \quad (-18.5) \times 100 / 43.5 = (-43\%) \text{ of unreliability.}\]

-That is to say the degree of reliability goes down of about 43% in the 3rd assessing attempt is concerned. Consequently, the unreliability reached 100% when counting the two rates.

\[
\text{57\% + 43\% = 100\% of unreliability}
\]

The unreliability is mainly recognised in the second and third assessment. In the second assessment try, the amount of points increases of 25 which means that the degree of reliability has lost about 57%. And, in the third assessment attempt, the amount of points increases of 18.5 which means that the degree of reliability has lost about 43%. If we count the unreliability of both assessments we can obtain an unreliability of 100%. That is to say, the change that has occurred in the second and third assessing procedures is an index of clear assessment inconsistency.
3.4. Teachers’ contributions (Multiple scorings)

The inconsistency of the obtained results increased the researcher’s doubt towards the legitimacy of his scoring procedure. The only way to clear up such a doubt was to ask for help from other teachers who are assumed to prove or disprove the existence of unreliability in assessing an essay-type test. Such unreliability is due to the teacher’s subjective interpretation of the essay answers. Is it possible to avoid subjectivity if we seek help of other teachers? Can we maintain consistent results as those obtained in scoring cloze test answers?

The three teachers were asked to assess the selected exam papers (See Appendix 8 and Appendix 9). The following table displays the teachers’ rates to the written essays samples.

<table>
<thead>
<tr>
<th>Written papers A/B</th>
<th>Teacher 1</th>
<th>Teacher 2</th>
<th>Teacher 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>11</td>
<td>09</td>
<td>10</td>
</tr>
<tr>
<td>B</td>
<td>10</td>
<td>07</td>
<td>8,5</td>
</tr>
</tbody>
</table>

Table 3.17. The teachers’ assessment to the selected written essays

The results of both papers are obviously different from those given by the threefold scoring attempts of the researcher. Such a difference can be recognised if we compare them in graphs.

The graph displays the unreliability of the assessment provided by the researcher and other teachers to the written paper sample A.

Graph 3.16. The unreliability of the multiple assessing procedures (Essay A)
The following graph displays the unreliability of the assessment provided by the researcher and other teachers to the written paper sample B.

Graph 3.17. The unreliability of the multiple assessing procedures (Essay B)

The two graphs refer to the written papers A and B. The numbers indicated in the bottom refers to the six assessing attempts. The numbers 1, 2, and 3 refers to the threefold attempts of the researcher; the others refer to the three scoring procedures of the enlisted teachers. As it is demonstrated in the graphs, the given marks of the same exam paper A and B are not on the same level. The scores, accordingly, change their weight from an assessing procedure to another. The essay was recognised by the researcher as being a failing work. The given marks seem inconsistent. Such inconsistency is also recognised in the other marks given by the three teachers, whose assessment does not only differ from that of the researcher in terms of the amount of points, they also increase the number of the passing works. That is to say, the researcher’s evaluation resulted in a number of failing and passing marks. The threefold attempts of the classroom teacher, for instance, resulted in 19 in the first assessing, 22 in the second and 19 in the third. The teachers’ results show that the written paper A is a passing work rather than a failing one. The first and the third teacher agreed to give it a passing mark, i.e., 11/20 and/or 10/20. The second teacher, even if he added 2, 5 to the same written paper, he regarded it as a failing work. The following diagrams may display better the unreliability
of the assessments provided by the researcher and the other three teachers who contributed to the study. The following graph displays the different marks allocated by the researcher and the other three participants teachers for the same written paper ‘A’.

**Graph 3.18. Essay “A” assessment unreliability**

The colours one, two and three represent the researcher’s assessments to the written essay A. On the other hand, the colours four, five and six correspond to the assessments of the first, second and third teachers respectively to the same written production A. The graph below displays the different marks allocated by the researcher and the other three participants (teachers) for the same written paper ‘B’.

**Graph 3.19. Essay “B” assessment unreliability**
The colours one, two and three represent the researcher’s assessments to the written essay B. On the other hand, the colours four, five and six correspond to the assessments of the first, second and third teachers respectively to the same written production B.

Such a disagreement over the assessment is also seen in the rating of the written paper B. The given marks are not the same as those given by the researcher. Each teacher allocated a completely different mark from the others. Such differences can be clearly recognised in the assessment of teacher, one who allocated a passing mark to the written paper B. Consequently, the number of the passing rates becomes 44 when adding two new written essays to the researcher’s first scoring. Both essays A and B were subjected to an unreliable assessing procedure. Such unreliability in assessment could be easily noticed in the following pie graphs. The graph calculated in a mathematically precise way the degree of unreliability in each assessment. The graph displays the different marks allocated by the researcher and the other three participants teachers for the same written paper ‘A’.

**Graph 3.20 Essay “A” assessment unreliability percentage**

In the previous pie graph, the colours one, two and three represent the degree of unreliability the researcher’s assessments to the written essay A. On the other hand, the colours four, five and six match the degree of unreliability of the assessments of the first, second and third teachers respectively to the same written production A. The following graph displays the
different marks allocated by the researcher and the other three teachers participants for the same written paper ‘B’

Graph 3.21 Essay “B” assessment unreliability percentage

In the same way, in the pie graph above, the colours one, two and three represent the degree of unreliability the researcher’s assessments to the written essay B. On the other hand, the colours four, five and six match the degree of unreliability of the assessments of the first, second and third teachers respectively to the same written production B.

3.5. The teachers’ interview

The three teachers who participated to the assessing procedure have been begged to answer as freely as possible the intended questions proposed in the interview (see Appendix 13). The interview was an attempt to decipher many sides of the issue of assessing essays. The teachers were asked about their attitudes and preferences towards the writing assessment process in particular.

The interview consists of six questions. The first important observation made about the teachers’ responses to the interview is that the three teachers had the same answers for the questions one, two and six. They all admitted that they prefer the essay rather than any other tool to assess their students’ written production. Moreover, they acknowledged that they used
the holistic approach to assess their students’ writing. More unexpectedly, the three teachers confessed that they did not provide any kind of written feedback to their learners’ essays since they holistically assessed such essays. As far as the other questions are considered here are their responses

**Question 3:**

**What does an essay at each of the following scales look like?**

**Teacher 1:**

1. **/ Excellent:**
   
The essay in this category contains a clear and direct answer to the question asked. The answer is developed through a sequence of reasonably ordered paragraphs. The essay is free from errors, and there will be evidence of superior control of language. In all regards, this essay is an excellent answer to the question, comprised of sharply focus, fluent, and interesting writing.

2. **/ Good:**
   
   In such an essay the student does not ignore the tasks of the assignment. The answer demonstrates a clear understanding of the writing task. It may not be as thoughtful or as carefully reasoned as the first essay, but it will not be characterised by mere statement and restatement of ideas at high level of generality. Although the essay may have minor weaknesses in paragraphing, it will contain evidence of the student's ability to organise information into fluent and unified paragraphs. The essay will be largely free from serious errors and will be generally well written, characterised by clarity, if not by sophisticated style.

3. **/ High average:**
   
The essay in this category will have an overall plan with a beginning, middle and end and will complete at least the basic tasks of the assignment. The reasoning may be somewhat imprecise and flawed, and the ideas may be repeated rather than explored. But it will be
organised and paragraphed well enough to allow the reader to move with relative ease through the discourse. It may contain errors, but not enough to continually distract the reader from the content.

4. / Low average:

The essay shows serious difficulty managing the tasks of the assignment. It will also fall in this category if it lacks an overall plan with a beginning, middle and end; if it shows definite weaknesses in development or illustration; if the ideas in paragraphs lack coherence; or if errors of mechanics seriously interfere with readability. Paragraphs may be somewhat disorganised, but the total effect will not be chaotic. The student's control of language may be imprecise, awkward, or clumsy.

5. / Weak:

The essay in this category may fail to come to terms with the assignment; that is, tasks may be ignored, misconstructed, badly mishandled, or redefined to accommodate what the student wants to say or is able to say. Such an essay may contain one or more of the following defects:

a- Serious errors in reasoning.
b- Little or no development of ideas.
c- Few or no connections between ideas.
d- Serious and frequent errors in mechanics may interfere with readability, giving the impression of distinctly inferior writing.

6. / Low / Very weak:

This category is reserved for the essay in which a combination of errors in English, conceptual confusion, and disorganisation create the impression of incompetence. Such an essay is typically marked by failure to focus on the main ideas, and recurrent grammatical and mechanical errors.
7. / Off-topic:
An essay in this category seriously interferes with comprehension. The student ignores the right answer. The essay is marred by serious errors in language form.

Teacher 2:
1. / Excellent:
The essay is well developed. It demonstrates no irrelevant or erroneous information. It leaves no doubt in the reader’s mind that the student possesses an excellent understanding of the question. Such a type of essay addresses both content and form in order to deserve such a grade.

2. / Good:
The essay is well developed. It may contain some erroneous information, but errors do not affect the overall quality of the essay. The reader must make some inferences because the answer is not always sufficiently explicit. Such an essay must address both content and form in order to deserve such a grade.

3. / High average:
The student basically understands the question, but the essay is not always well-focused. It may contain significant errors of fact or interpretation. If the essay addresses only content, the treatment must be very good to deserve such a grade.

4. / Low average:
Poor organised essay. Limited understanding of the question, and may contain major errors or to be so general as to suggest that the student is unable to deal completely with the question.

5. / Weak:
The essay is chaotic, confused and incorrect. The reader is left with the certainty that the student has not clearly understood the question.
6. / Low/ very week:
The answer is on the task, but is so brief or so poorly written as to be meaningless.

7. / Off-topic:
The essay is off-task. The student basically misinterpreted the question, although the essay shows a remarkable writing ability.

Teacher 3:

1. / Excellent:
The essay is fluent, well-developed, and well organised. It is an excellent essay. It completes all the tasks set by the assignment. It distinguished by the quality of its content and the amount of detail used to support or illustrate central ideas. The student clearly demonstrates a superior control of language and sentence structure. Such an essay avoids all types of errors in grammar, and mechanics.

2. / Good:
Such an essay demonstrates all the tasks set by the assignment, but it may be less developed or less fluent than the first essay. The student clearly understands the question, provides sufficient detail to support central ideas, and demonstrates an ability to organise information coherently and use language effectively. Although it may have a few minor errors, such an essay is largely free of serious errors in mechanics and grammar.

3. / High average:
The essay is an adequate essay. It may slight one of the tasks of the assignment or deal with it only by implication, but it demonstrates a clear understanding of the topic. Such an essay tends to be less well developed than the second essay, but it shows a basic competence in its handling of organisation, sentence structure and grammar. There may be some awkwardness or disjointedness, an imprecise use of language, or minor errors in mechanics, grammar, usage
and sentence structure, but the writing problems will not be serious or frequent enough to
distract or confuse the reader.

4. / Low average:
The essay is an inadequate university level writing. It usually lacks sufficient development or
the writer’s commands of language appear uncertain . It typically suffers from one or more of
the following weaknesses:
a-   Definite weakness in defining the subject-matter.
b-   Major problems in organisation, focus and coherence.
c-   Insufficient specific information to illustrate or support arguments.
d-   Serious problems in grammar, usage, and mechanics that are frequent and serious enough
to distract or confuse the reader.

5. / Weak:
The essay is a weak work. It typically shows an inability to answer directly the question. In
addition to problems in development and organisation, such an essay will usually be marked
by problems in sentence coherence, usage and grammar.

6. / Very weak:
The essay is almost completely unsatisfactory. It is so slight that the student appears to have
been overwhelmed by the assignment, or so error-hidden that it is incoherent.

7. / Off-topic:
Such an essay ignores the terms of the assignment. The teacher may award an extra-point to
exceptionally well-written papers that fail to deal fully with the topic. This should be done
only on those occasions when the writing is much better than the content.

It is obvious enough from the teachers’ answers that there is no agreement among them
about the criteria or features of each of the essay’s levels. Such a fact may automatically
affect their assessment process. This may explain their disagreements upon the marks they allocated to the selected written pieces.

**Question 4:**

Which marks do you prefer to provide to the following essays’ scale?

**Teacher 1:** The table displays the teachers’ rates to the different assessing level.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Marks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. /Excellent</td>
<td>13/20</td>
</tr>
<tr>
<td>2. /Good</td>
<td>12,5/20 - 12/20</td>
</tr>
<tr>
<td>3. / High average</td>
<td>11,5/20 - 11/20</td>
</tr>
<tr>
<td>4. / Low average</td>
<td>10,5/20 - 09,5/20</td>
</tr>
<tr>
<td>5. / Weak</td>
<td>09/20 - 08/20</td>
</tr>
<tr>
<td>6. / Low/ very weak</td>
<td>07,5/20 - 06/20</td>
</tr>
<tr>
<td>7. / Off-topic</td>
<td>05/20</td>
</tr>
</tbody>
</table>

*Table 3.18. Teacher 1 rates to each an assessment scale*

**Teacher 2:** The table displays the teachers’ rates to the different assessing level.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Marks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. /Excellent</td>
<td>14,5/20</td>
</tr>
<tr>
<td>2. /Good</td>
<td>13,5/20 - 12/20</td>
</tr>
<tr>
<td>3. / High average</td>
<td>11,5/20 - 10/20</td>
</tr>
<tr>
<td>4. / Low average</td>
<td>09,5/20 - 08,5/20</td>
</tr>
<tr>
<td>5. / Weak</td>
<td>07,5/20 - 06/20</td>
</tr>
<tr>
<td>6. / Low/ very weak</td>
<td>05/20 - 04/20</td>
</tr>
<tr>
<td>7. / Off-topic</td>
<td>03/20 - 5/20</td>
</tr>
</tbody>
</table>

*Table 3.19. Teacher 2 rates to each an assessment scale*
Teacher 3: The table displays the teachers’ rates to the different assessing level.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Marks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Excellent</td>
<td>12/20</td>
</tr>
<tr>
<td>2. Good :</td>
<td>11.5/20 - 11/20</td>
</tr>
<tr>
<td>3. High average :</td>
<td>10/20 - 09/20</td>
</tr>
<tr>
<td>4. Low average :</td>
<td>08.5/20 - 07.5/20</td>
</tr>
<tr>
<td>5. Weak :</td>
<td>07.5/20 - 06.5/20</td>
</tr>
<tr>
<td>6. Low/ very weak :</td>
<td>05/20 - 03.5/20</td>
</tr>
<tr>
<td>7. Off-topic :</td>
<td>02/20</td>
</tr>
</tbody>
</table>

Table 3.20. Teacher 3 rates to each assessment scale

The answers to the question three have revealed the teachers’ disagreements about the characteristics of each essay level. Consequently, such disagreements affect unquestionably also the marks they are likely to allocate to such essays. Their answers to question five came to confirm their disagreements. An off-topic essay may be allocated 02/20 according to the third teacher whereas the two other teachers mentioned 05/20 for the same essay level. In the same way, the first teacher might allocate 13/20 for an excellent essay; the second teacher spoke about 14.5 for the same essay, while the third teacher pointed out to 12/20.

On the other hand, two teachers agreed upon the mark to be given to a weak essay. Both of them mentioned 7.5 to 6.5/20. They also agree about a very weak essay since they might allocate 04 to 05/20 to such an essay.

All in all, teachers’ marks to each written essay scale seem also unreliable and inconsistent.

Question 5:

Which of the following essay components do you prefer to rely on in your assessing procedure?
The teachers’ answers to the previous questions are summarised in the following table which displays the written components that the teachers consider in their assessment.

<table>
<thead>
<tr>
<th>Teacher 1</th>
<th>Teacher 2</th>
<th>Teacher 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>-Content: interest/ originality of ideas</td>
<td>-Content: interest/ originality of ideas</td>
<td>-Content: interest/ originality of ideas</td>
</tr>
<tr>
<td>-Organisation of the ideas</td>
<td>-Organisation of the ideas</td>
<td>-Organisation of the ideas</td>
</tr>
<tr>
<td>-Rhetorical matters : organisation, cohesion, unity</td>
<td>-Rhetorical matters : organisation, cohesion, unity</td>
<td>-Rhetorical matters</td>
</tr>
<tr>
<td>-Relevance of the content</td>
<td>-Relevance of the content</td>
<td>-Relevance of the content</td>
</tr>
<tr>
<td>-Mechanics</td>
<td>-Diction( word choice)</td>
<td>-Grammar/ Syntax</td>
</tr>
<tr>
<td>-Diction( word choice)</td>
<td></td>
<td>-Mechanics</td>
</tr>
</tbody>
</table>

Table 3.21. Teachers’ assessed written components

As far as the last question is concerned, however, teachers have shown a high sense of agreement upon which components to consider in the assessment process of any written production essay. They all agreed to consider four significant components: (1) content: interest/ originality of ideas, (2) organisation of the ideas, (3) rhetorical matters such as organisation, cohesion, and coherence, and (4) relevance of the content. The disagreement was just about mechanics, diction and grammar as displayed in the Table 3.19. More surprisingly no teacher mentioned the neatness of the paper, the form: effectiveness of expression, the handwriting, or the style as major components in any assessing procedure of a written essay.

To sum up the researcher deduced a kind of a paradox. The teachers agreed upon which components to consider in the assessment of an essay. On the other hand, they were entirely in disagreement as far as rating an essay is concerned.
3.6. The results’ interpretation

The present modest study has confirmed that using the essay as an assessing tool is one of the main controversial issues that is still a matter of examination and enquiry. The assessing formula is most of the time blamed to be holistic and entirely subjective due to the fact that it is not an easy task to understand and identify the bases upon which the rates would be provided. The teacher’s judgment is rather shaped by individual views and beliefs, and prejudices which may provoke untrustworthy marks. Such an issue is still a challenging problematic in testing and assessing the language writing performance.

On the other hand, students’ writing as a response to objective tests such as direct questions or cloze test items have never been concerned with such problematic. The answers are most of the time built upon specific information. Thus, the assessment is usually an easy task, and the marks are often correct in a scientifically precise way. The teacher’s assessment is, to a great extent, based upon a countable process. No judgement is needed from the part of the classroom teacher or any other assessor who may be invited to provide his estimation to the student’s writing performance. The marks allotted remain unchangeable from an essay to another, and from an evaluating settings to another. This is due to the fact that the students’ answers are simply either right or wrong.

Consequently, the use of the two different written tests (the essay and the close-test) with both second and third year EFL learners at the University of Sidi-bel-Abbes was an attempt to check up whether the EFL teachers’ assessment to students’ writing is reliable and objective or not. Simply put, the study attempted to determine the extent to which the teachers’ assessment to students’ writing could be reliable and valid or not.

The findings show that reaching a high level of reliability in assessing the learners’ written production or the essay in particular is still a far reaching objective. The assessment of the first written test suggested by the researcher has shown a kind of high reliability. Such
reliability was simply due to the test’s nature as being a cloze test. However and more amazingly the reliability in assessing such cloze test was seen only in the passing and failing rates which remain static from one assessing procedure to another. In other words, when checking the allocated marks, a sort of unreliability was revealed. The study, thus, pointed out the reliability of passing and failing rates and the unreliability of the assessment procedure.

However, it is tremendously significant to mention that such kind of written tests (mainly the cloze test, the MCQ test, matching test, true or false test, reordering test and so on) do not tell the teacher a lot about how well his students are improving their abilities and competences. Therefore, he could never design an effective feedback to address the situation. As far as the researcher view is concerned, such kinds of tests are neither valid nor reliable. Students often provide answers to an MCQ test, for example, without making great effort to think about the content of the question. The trouble is that several times the students’ answers coincide with the right ones; therefore a good mark would be allocated. And sadly, such mark would reflect neither the student’s strengths nor his weaknesses. It would simply mislead both of the teacher and the learner.

On the other hand, the teachers’ assessments to the second written test or the essay administrated to the third year students revealed a high level of unreliability. The researcher with the help of three other colleagues had tried hard to achieve reliability in assessing the students’ essays, without any amazing success. To start his survey, the researcher assessed holistically ninety-two written papers in three different occasions. However, none of the mark given to each written essay resembled to another.

More unexpectedly, the assessments provided by the teachers who were begged to participate to the study reached a high degree of unreliability. Despite the fact that those three teachers were invited to assess just two selected papers ‘A’ and ‘B’, they have never come to an agreement about the mark to be allocated. No one can dare to deny that assessing an essay
holistically may be unjust to students especially when it comes to tests or exams upon which their study success and continuity are dependant. Moreover, and even worse, holistic assessment procedure is not so helpful since it would never reflect how well students’ are progressing. Thus, the assessment reliability and objectiveness are not attained. The teacher would never recognise his learners’ deficits and troubles with the language. Therefore, he would never have such opportunity to provide them with appropriate feedback to cure and repair the lacunae. The learners, on their part, would never been exposed to their mistakes, so they would stay indifferent and unresponsive vis-à-vis such impediments and hindrances. They would never have the chance to auto-correction and auto-evaluation. In holistic assessment the compositions are read for general impression and according to this impression the teacher judges his students’ level of proficiency. All the aspects of the composition—content and conventions affect the teacher’s response, but none of them is specifically identified or directly addressed using a checklist. Teachers can never chart students’ development and gather information that will help them determine grade and quality. This approach is rapid and efficient in judging overall performance. It may, however, be inappropriate for judging how well students have applied specific criterion or developed a particular form.

The analytic assessment, however, considers writing to be made up of various features, such as relevance, paragraphing, style, organization of ideas, expression of thoughts, grammar and mechanics each of which is to be assessed separately. The composed essays are read and assessed through the focus on a pre-determined list of criteria. Such criteria can be compared to a set of standards. The teachers use consistent assessing scheme based on selected criteria of evaluation in order to assess each composition aspect separately. Doing so, the assessment procedure would ensure reliability and validity. Besides, an effective feedback would be the outcome of such assessment. Moreover, the teacher would draw many conclusions about his
students’ writing performance. He would also keep a full profile of his learners’ writing aptitudes in mind. It also allows the learners to identify areas in their essays that need extra effort when accompanied by written comment and remedy. The analytic assessment diagnostic nature provides students with a road map for improvement.

Finally, the teachers’ interview has divulged many hidden truths. The teachers were always on disagreement about the ingredients of an excellent, average or weak essay. Moreover, they could not share the same view about how much merit should be assigned to each essay scale. Besides, they disagreed about the components to be considered when assessing an essay. Therefore, the unreliability of the assessment they had provided to the students’ writing was a logic outcome of the different views they share about the assessing procedure. In this respect, the computer would achieve a better reliability and objectivity of the assessing process, since everything would be calculated in mathematically precise way. Moreover, the computer holds no prejudices, no feelings, no own views, and no preferences. It would do the work as neutral as possible. Waiting for the dreams to be a reality, writing teachers would coordinate and cooperate to unify views about assessing writing.

A valid test reliably assessed would provide an effective feedback that no doubt would enhance the students’ writing performance. Research has shown that the learners’ writing improvement requires two important actions from the part of the writing teacher: First, he should expose the learners to their immediate lacunae. In other words, he ought to make the learners recognize and identify not only the mistakes, but also their nature (relevance, style, diction, grammar, mechanic, paragraphing or coherence). Second, an effective teacher might provide his students with sufficient and efficient tools to correct not only the mistakes they commit, but also the mistakes they will commit. Corrective sessions are so useful and effective as it might be proved in the next chapter.
3.7. Conclusion

The traditional assessment of EFL learners’ writing still has a legitimate place in the English language arts, but could not be the secure means of assessing writing. Evaluation is vital for a clear, reliable picture of how students are progressing and how well the methods of instructions address their needs. The assessment of written language can take two forms of procedures: the holistic and analytic assessment.

As previously mentioned, the writing teachers usually go through two different writing tests the objective and subjective ones. The assessment of an objective test is usually simple and reliable to a great extent, as it was the case with the assessment of the secondary population of the study (the second year EFL learners). However, assessing an essay is very hard task that always result in the unreliability of the assessing process and in the effectiveness of the teachers’ feedback, as it was the case with the written test of primary population of the study (the third year EFL learners). Such unreliability is clearly noticeable in the dissimilar marks allocated for the same papers during three different occasions by the same assessor.

The analytic assessing method has proved to be the most effective method that can, to some extents, reach conclusive and good results. This may be because it is based on a psychometric/atomistic approach. Such a method can arrive at suitable results if they are effectively used. An effective assessing procedure needs an increasing concern from the part of the classroom teacher who is assumed to be in charge of the construction and the administration of the composition topic. If the construction of a topic needs some steps such as planning the composition, preparing the composition items and directions, the assessing procedure must also be based on an organized technique that could be carried out in order to achieve improvement. The effectiveness of assessment is highly dependent upon the manner in which the procedure is employed.
In general, the method that could be accepted is the analytical one. The weaknesses can never be obstacles to maintain the credibility of such a method. Experiences have shown that the analytic method can be formally used by the classroom teacher, thus the pupils’ behaviour in writing is easily controlled. The assessment systems as well as evaluation are closely tied to a teacher’s own personal philosophy regarding the best way of assessing.

However, the analytic procedure is usually blamed by teachers to be time and effort consuming. The teachers are usually complaining about the enormous number of students’ papers that they have to assess and correct. Assessing analytically an essay means evaluating every single component alone. Therefore, great deal of time and effort is required to achieve an effective assessing process.

Trough an analytic assessment, EFL learners would no doubt acquire new behaviours when dealing with the writing activities. They would understand that any written expression needs an introduction, a development, and a conclusion. They would become aware that the lack of coherence and cohesion means confusion, and the absence of punctuation means run-on sentences, fragments, and confusions in meaning. The students’ failure in writing is mostly due to their unawareness of their errors and thus their inability to get rid of them, rather than to their communicative and linguistic competencies.
Chapter Four

Pedagogical Implications
# Chapter Four

## Pedagogical Implications

<table>
<thead>
<tr>
<th>Section</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1. Introduction</td>
</tr>
<tr>
<td>4.2. The Formal Assessing Procedure</td>
</tr>
<tr>
<td>4.2.1. The Selection of the Assessing Components</td>
</tr>
<tr>
<td>4.2.2. The Specification of the Criteria of Assessment</td>
</tr>
<tr>
<td>4.2.2.1. The Relevance</td>
</tr>
<tr>
<td>4.2.2.2. Paragraphing</td>
</tr>
<tr>
<td>4.2.2.3. Written Expression</td>
</tr>
<tr>
<td>4.2.2.4. Organisation and Coherence</td>
</tr>
<tr>
<td>4.2.2.5. Grammar and Mechanics</td>
</tr>
<tr>
<td>4.2.3. The Architecture of an Analytic Assessing Scheme</td>
</tr>
<tr>
<td>4.3. The Macrostructural Assessing Procedure</td>
</tr>
<tr>
<td>4.4. The Microstructural Assessing Procedure</td>
</tr>
<tr>
<td>4.4.1. Assessing the Paragraphing of Ideas</td>
</tr>
<tr>
<td>4.4.2. Assessing the Organisation and Coherence of Ideas:</td>
</tr>
<tr>
<td>4.4.3. Assessing Language Form</td>
</tr>
<tr>
<td>4.4.4. Assessing the Style</td>
</tr>
<tr>
<td>4.5. The Analytic Assessing Limitations</td>
</tr>
<tr>
<td>4.6. Automatic Essay Assessment</td>
</tr>
<tr>
<td>4.6.1. Project Essay Grader (PEG)</td>
</tr>
<tr>
<td>4.6.2. Intelligent Essay Assessor</td>
</tr>
<tr>
<td>4.6.3. E-Rater</td>
</tr>
<tr>
<td>4.6.4. IntelliMetric</td>
</tr>
<tr>
<td>4.6.5. Bayesian Essay Test Scoring System</td>
</tr>
<tr>
<td>4.7. Designing an Effective Feedback</td>
</tr>
</tbody>
</table>
4.8. Reshaping the Writing Feedback

4.8.1. Where Am I Going?

4.8.2. How Am I Going?

4.8.3. Where to Next?

4.9. Enhancing the Effectiveness of the Feedback.

4.9.1. Reflecting Upon the Feedback Process

4.9.2. Considering Students’ Needs and Suggestions

4.9.3. Varying Feedback Brands

4.9.4. Feeding Up

4.9.5. Equilibrating Positive and Negative Feedback

4.9.6. Feeding Forward

4.9.7. Incorporating Remedial Task

4.10. Planning a Corrective Conference

4.11. Managing the Corrective Conference

4.11.1. The Review of the Essay Question

4.11.2. Teacher’s Impressions

4.11.3. Collective Correction

4.11.4. Individual Correction

4.11.5. Remedial Work

4.12. Conclusion
4.1. Introduction

It is often conventionally assumed that tests are mostly used for assessment: the test receives a mark upon which the learners’ level of proficiency is defined. Such tests or rather their assessment can provide the teacher with information about his students’ achievement of the instructional goals. They can also reflect the degree of success not of individuals but the instructional program itself. Teachers are assumed to prepare the tests in the same way they do with the courses. There are now a number of very effective methods of teaching English as a second or foreign language. There also many testing and assessing methods which are mostly used for objective tests.

However, the area of testing and assessing is still lacking researches and investigations mainly when it comes to assessing subjective tests such as the essay. The present chapter tries to stress the usefulness of two main components to enhance the effectiveness of writing assessment mainly the use of a highly designed analytic assessment method through a highly designed computing programme.

The analytic assessing approach has proved to be the most objective and effective method that can to some extent reach a conclusive and reliable assessment. This may be because it is based on a psychometric/atomistic approach. Such a method can achieve appropriate results if it is effectively used. A formal assessment procedure needs an increasing concern from the part of the classroom teacher who is assumed to be in charge of the construction and the administration of the test question. If the construction of an essay question needs some steps such as planning the test, preparing the test items and directions; if the administration of such a question includes fairly detailed instructions in order to validate the test process, the assessing procedure must also be based on an organised techniques that should carried out in order to achieve reliable results. The reliability of the test assessment is highly dependent upon the manner in which the assessing procedure is employed. The assessing procedure,
accordingly, is based on four considerable steps: the selection of the assessing criteria, the specification of the criteria of evaluation, the specification of a rating plan and the selection of the test question. Such steps are organised in a way that permit the classroom teacher to be more objective in his judgements of the students’ responses.

Computer, in this respect, seems to do thing much better than any human teacher. Innumerable are the researchers that have examined the reliability extents of the computer assessments such as Burstein & Chodorow in 1999 and Landauer, Laham, & Foltz, in 2003. The findings have revealed an important agreement rates between the machine and the human. Automatic assessment is possible today thanks to text analysis software. Such an assessment may consider not the overall quality of the text but also certain features. The effectiveness of the assessing software was demonstrated in three application areas: the accurate answers to open questions, essays’ assessment, and the detection of plagiarism. The researcher suggests an analytical assessing scheme which may be developed into computing software towards the achievement of reliability in the assessment of the EFL learners’ writing.

4.2. The formal assessing procedure

Prior to any challenge to react to students’ written production, there are steps that the teacher has to go through principally the selection of the assessing components, and the criteria for assessment. Such steps aim at boosting the whole analytic assessment process.

4.2.1. The selection of the assessing components

An essay is a type of test in which the learner is asked to respond to the essay question set by the assessor. The answer should consist of all the essay components. The learners are to be assessed not only on the straight reply, but also on their employ of all the components which combine an essay. The significance of the key items such as discuss, comment, explain, compare, analyse, enumerate, evaluate, state, etc are utilized to direct the learner’s attention to the kind of paper they should deal with.
The aptitude to write an essay engages many unified and consistent components. These components guide to obvious, fluent, and effective communication of thoughts. Most teachers and specialists in assessing and testing English would almost certainly agree about categorizing a number of varied elements which compose the fundamental supports of the writing process. Such fundamentals may differ from a researcher to another, but they all agree on the most central of them. Consequently, Lado (1962) sees:

There are things that can measured in connection with content the points of information to brought out; the organisation and sequence in which these points are presented; the formal signals given the reader to guide him in understanding the topic fully (Lado, R; 1962: 248)

The writing course as it is assumed by Lado comprises three major components: the relevance of the content, the organisation of ideas and the paragraphing. These components are imperative but there are also two other crucial ones: Language accuracy and style. More progress would be made, according to him, by assessing language as language, and content and style as content and style. Therefore, in an essay, five major components should be assessed: relevance, organisation of ideas, paragraphing, language and style.

Besides, Harris (1969: 68-69) believes that the writing process is to comprise five components: First, the content or the substance of the writing (the expressed ideas), second; the form or the organisation of the content, third; Grammar, fourth; the style or the selection of configurations (structures) and of lexical items to set up a meticulous quality or taste to the writing, and finally the mechanics or the use of the graphic conventions of the language: spelling, punctuation, and capitalization.

It is understandable that Harris & Lado stress language and style as key components of the writing. Language includes both grammar and mechanics, and style includes diction and writing fluency. The form, however, refers to the organisation of ideas in an essay. According
to Lado & Harris (1962) the writing process includes six items: First the content; second the organisation of ideas; third the paragraphing; fourth the grammar; fifth the mechanics and sixth the style. These six components have been probably recognised as the most important elements of an essay.

In her turn, Raimes (1981) sees that the writing process unites many consistent components. He categories them as content, organisation, grammar, syntax, mechanics, word choice, purpose, audience, and the writer’s process. According to him, all these components show the way to clear, fluent, and effective communication of ideas. As opposed to Lado (1962) & Harris (1969), Raimes (1981) inserts three potential components: First, the rationale of writing, second; the audience and third, the writer’s process. Such components are mostly used in writing a free-composition. A composition test permits the leaner to write his own moderately free and complete reply to the topic set by the teacher, (Harris; 1969: 05). The essay test, on the other hand, requires the learner to react to the question-matter in an essay framework. The two different tests share the same six components cited above. The main difference is made when the purpose of the test is set forth.

Yorkey (1982) has also tried to determine the components to be considered in any written piece assessment.

An essay-type examination is verbal communication. The clarity of the message depends upon the clarity of your expression. If your grammar is imprecise, if your vocabulary is ambiguous, if your organization is distorted, if your handwriting is illegible, there is likely to be a breakdown in communication. Even if the message comes through, confused but comprehensible, your teacher may unconsciously deduct for straining their eyesight and patience. (Yorkey; 1982: 236)

According to Yorkey the writing process should include six components: First, the message, second, the style (written expression), third; grammar, fourth; vocabulary, fifth; organisation,
and sixth; handwriting. **Yorker** regards handwriting as a fundamental ingredient of the writing process. The lucidity of the communication, according to him, is primarily dependent upon the clearness of the handwriting. It is, therefore, an indispensable component that should be given a special emphasis.

**Dickens and Germaine (1993: 51)** have also investigated the issue. They admit the existence of twelve components that can be included in any writing process. Such components are of two types. The first type includes those that centre on the accuracy of language use, such as grammar, vocabulary and tenses. The second includes those that focus on communication such as style, appropriateness, attempt to communicate, fluency, and relevance or significance of the content.

Moreover, the purpose of writing, according to **Ur (1996: 163)**, is the expression of ideas, the transmission of a message to the reader, so the ideas themselves should debatably be seen as the most facet of the writing. On the other hand, the writer is also invited to pay more awareness to formal characteristics: tidy handwriting, accurate spelling and punctuation, as well as satisfactory grammar and vigilant selection of vocabulary. In such a condition, the essay consists of two interconnected pieces: language form and content. The former includes spelling, grammar, punctuation and handwriting. The latter, on the other hand, includes interest, originality of ideas, effectiveness of expression and organisation.

Finally, **Kenji. K. & S. Kathleen K. (1999)** see the writing process as a highly complicated skill combining an amount of various basics, only a number of which is firmly linguistic. The ability to write involves at least six components. First, the grammatical ability or the capability to write English in grammatically accurate way, second; the Lexical aptitude or the skills to appropriately choose and use the words, third; the mechanical skill or the ability to perfectly use punctuation, spelling, capitalisation, etc., fourth; the stylistic skills or the talent to use sentences and paragraphs appropriately, fifth; the organizational
skills or the ability to systematize the written work according to the conventions of English, including the order and range of material and finally the judgements of appropriateness or the ability to make judgements about what appropriate depending on the task, the purpose of the writing, and the audience.

The preceding review shows clearly the intricacy of the writing skill. Such a mission, as usually visualized, is a well-built tool which would necessitate additional concentration and a cautious measurement of its components. Apparently, these components comprise the key standards of assessment that the assessor would take consider in order to maintain a reliable assessing process. It would be of a paramount significance to decide on in advance what is to be assessed in an essay. According to Dickens and Germaine; teachers should in advance of administering a test which criteria they will use to mark learner’s work. (Dickens & Germaine; 1993: 50). A careful range, therefore, appear to guide to an enhanced understanding of the major point of an essay test. The essay is a sort of oral message which requests all the properties which form its primary basis.

The choice of the components of assessment, as it is supported by Mahili (1994), is the most essential detail that the assessor has to take into account prior to the beginning of the assessing procedure. According to him; some teachers tend to impose themselves as authorities and make comments reflecting the application of an ideal standard rather than having a set of criteria for marking. (Mahili, I.; 1994: 24)

Accordingly, the criteria of evaluation have immense impacts on rates. It has been found that among the disadvantages of the holistic assessing process is the lack, or the unreliability of the criteria of assessment. The uncertainty in choosing the assessing conditions would have unhelpful consequences on the assessing procedure. Educators may find themselves powerless to validate and even to manage such an essay assessment.
In any essay there are countless components that can be assessed. The assessor ought to border the amount of components in order to direct in a very accurate method his assessing process. According to Madsen (1983);

One reason to evaluate a few factors at one time is that doing so helps us grade our papers more accurately and consistency. Another reason is to speed up our essay grading. A third reason for limiting the number of factors to be evaluated is to avoid unnecessary discouragement of our students.

( Madsen, H.S. ; 1983:119 )

The selection of the chief features that can be weighed up is a vital element in language assessment. It may allow the teacher to conduct his assessing procedure in a further objective mode. Such a choice may have three optimistic tenets. First, it assists the assessor to synchronise his judgements to some definite components. Second, the specification of the criteria of assessment can speed up the assessing procedure, and makes it easy and practical. Third, such exactitude in determining the number of components may have a positive outcome on the students' attitudes towards the issue of assessment.

The criteria are usually of two brands. The first brand contains all those criteria that focus on the correctness of language use, such as grammar, mechanics, and diction. The second brand, however, has been highlighted by communicative approaches to language teaching, and includes criteria such as the relevance of content, organisation of ideas, paragraphing, and style. These two types of criteria constitute the basis of any scoring procedure.

4.2.2. The specification of the criteria of assessment

Considering all the previous view about the constituents to be assessed in any EFL learner’s writing, it can be asserted that there are seven criteria that should be taken as the main bases of any written piece assessment. Each of them is divided into categories or levels that may rank the essay. So in order to go through a reliable and objective assessment that
would result in a sufficient and efficient feedback, it is advisable to consider and to examine each of the following components:

4.2.2.1. The relevance

The relevance of an essay refer in the information and data suggested by the learner as an answer to the essay question. Such relevance may appear in four different levels according to the savoir-faire and the aptitude of the learner.

To start with, the written production may demonstrate superiority. It leads no doubt in the reader’s mind that the student possesses a higher perception of the question. The essay would entirely and straight answer the question. It holds up the answer with thorough proof. It offers insights not willingly obtained from class lectures. All the suggested information are right. Second, the essay may minimally responds to the question. It shows facts of serious reflection of the theme. It supplies a number of supporting data derived from the teacher’s courses. Some information are correct. Third, the essay almost replies to all or most part of the enquiry. Its supporting information are not enough, deficient, or incorrect. A part of the answer is mistaken. Fourth, the essay answers less than half of the subject. It suggests no supporting evidence. Most of the answer is incorrect. The essay is a failing work. There is a misconception of the subject. The reply is completely off-topic. The reader is left with conviction that the student has not understood the question.

4.2.2.2. Paragraphing

Paragraphing refers to the overall structure of the designed paragraphs. Three levels may be identified. First, the essay may enclose an obvious, brief and straight response to the topic and widens that response through a succession of convincingly well-organized paragraphs. The essay has a detectable beginning, middle, and end. The introduction focuses the reader’s awareness on the topic of the essay in a careful paragraph of thought, suggesting sentences leading successfully into the thesis statement which evidently, particularly, and captivatingly state or involve the key thought or thoughts which the essay would make clear or support.
The development or the body meticulously sustains and clarifies the thesis or builds to a reasonable end in a sequence of glowing, attractive paragraphs. And a conclusion logically accomplishes the expansion of the thesis or build up to the main points of the essay.

Second, the essay has slight flaws in paragraphing, but it includes confirmations of the writer’s skill to arrange data into smooth and amalgamated paragraphs. The paragraphing is limited in logical development of ideas into paragraphs. Both introduction and conclusion are lacking.

Third, significance complexity is observed in controlling the assignments. The essay is short of an overall plan with a beginning, middle and end. The paragraphs are somewhat jumbled. The key ideas in paragraphs lacks development or demonstration. They offer modest evidence of the student’s skill to build up a prearranged essay.

4.2.2.3. Written expression

The essay could be assessed at three different levels. First, the essay demonstrates mastery of the elements of effective writing. It displays an effective range of sentence clarity, variety and word-choice. Second, the essay demonstrates sentence clarity, variety and generally appropriate choice. Some inappropriate words do not affect the quality of the essay. The meaning is comprehensible. Finally the essay demonstrates virtually no sentence variety. The meaning of the sentence is ambiguous due to the consistently inappropriate or non-idiomatic word –choice. The writer’s control of language may imprecise, awkward, or clumsy.

4.2.2.4. Organisation and coherence

Organisation and coherence are to important features of a well written essay. Three scales may be noticed in the learners’ essays. First, the ideas may persuasively be developed. They are reasonably planned and linked with obvious transitions. The writer displays patent understanding of the writing assignment. He has a superior control over the language to
express thoughts with practical excellence. Second, the ideas are agreeably progressed and ordered, but they are not linked with transitions. Third, the essay is disorganized, puzzled and mismanaged. There are no relations between ideas.

4.2.2.5. Grammar and mechanics

Grammar and mechanics polish the written production since their misuse may cause many trouble to grasp the message. The essay may demonstrate five grammar and mechanics scales. The essay may first reveal reliable mastery of language and mechanics. It is practically without faults of grammar and mechanics, and there is a kind of finer management of the language. Second, the essay expresses proficiency. It is free from grave errors and is usually well-written and characterised by transparency. Third, the essay encloses some errors in grammar and mechanics, but not to repeatedly distract the reader from the context. Fourth, the essay is short of competence. It exhibits severe and numerous troubles in the use of language and sentence construction. It contains frequent errors in grammar, usage and mechanics that impede the meaning. Such an essay signals that the learner is unable to deal proficiently with the subject. Fifth, the essay makes obvious a whole deficiency of skill. It has harsh and constant errors in language and sentence structure. It holds a persistent prototype of errors in grammar usage or mechanics that may interfere with readability giving the feeling of definitely poorer writing.

4.2.3. The architecture of an analytic assessing scheme

The third measure that the teacher can take is to set up an analytic scheme based on the particular criteria of assessment. He should make a decision in advance on the specific bases for assessing. To Harrison (1990), the assessing system, according should be carefully studied before any attempt to test students’ abilities. According to him;

The marking scheme should be thought out at an early stage in the development of the test, since it is in principle a forecast of
what the students will produce and so affects what is to be included in the assessment. (Harrison. A. A; 1990: 111)

Our assessments to the students’ responses, whether they are formal or informal correct or mistaken, are mostly based on individual views. It is relatively probable to make such assessment sound reliable and objective. The opportunity to attain such an objective is to detail the criteria of assessment in an assessing framework and should be tried out on a experiment version so that any required assessments can be made. This assessing system, as it is recommended by Basanta (1995), should be checked up at least by another teacher. According to him; the marking criteria should be set before hand and candidates must be informed as how they will be scored. (Basanta. P. C.; 1995: 57)

Students are supposed to recognize how they will be assessed. Such an intention may permit them to be more aware of the type of test they are supposed to respond to. Each student is able to see how his particular grade has been obtained (Heaton. J.B; 1975: 137). He is in a situation to assess the quality of his work by himself. Moreover, each student will be optimistic to adapt his essay to the task of the assignment without a panic of being penalised for something he does not know.

The design of an analytic assessing scheme can be also sustained by a lucid explanation of its constituents. Each constituent must be studied and controlled through the determination of its qualities that it is based on. Consequently, the analytic assessing procedure is to consist of assessing in categories such as the relevance of content, paragraphing, organisation of ideas, diction, grammar, mechanics and style assigning each a mark and awarding total mark out of twenty as it is displayed in the researcher’s own adapted analytic assessing process diagram. The following diagram illustrates the essays’ analytic assessment as suggested by the researcher. However the points allocated to each component could be modified according to the immediate objective of the writing class.
It is probable to use the Process Approach (Mahili, 1994) which engages manifold drafts focusing on both content and language at different separate phases. Such an assessing process is made up of two procedures: Macrostructural and Microstructural.

4.3. The macrostructural assessing procedure

As its name may indicate, the Macrostructural procedure considers the broad assessment of the relevance of content. It scans and examines the general idea of the essay without any orientation towards the other criteria of assessment such as paragraphing, organisation, diction, grammar, mechanics and style which on the other hand refer to the main concern of the Microstructural procedure. The teacher generates an answer sample to the subject-matter in order to be used as a reference when assessing the learners’ essays. The answer is a set of ideas, and each idea is allotted a point. The amount of points constitutes the general mark allotted to the relevance of content. The following figure represents the answer sample as suggested by the researcher to the written test provided for the main population of the present
study namely third year students. The answer is based upon the teacher’s knowledge about the subject matter.

**The essay answer sample**

- **The essay question**

“As opposed to Microlinguistics, Sociolinguistics has been established as a multidisciplinary field of investigation with a distinct approach.” Discuss.  

**The answer is based on two main ideas: 06/06**

1. Microlinguistics as an independent study versus Sociolinguistics as a multidisciplinary field of investigation.

2. Sociolinguistics adopts a distinct approach which is quite different from microlinguistic approaches.  

**The suggested answer sample**

1. Microlinguistics is seen as an independent discipline. It studied language without any reference to other approaches. It set up its own methods which did not share with the other disciplines. (01pt)

2. Sociolinguistics as a new field of investigation is mainly dependent upon four disciplines such as Generative Grammar, Dialectology, Anthropology, and Sociology. (01pt)

3. **Microlinguistic approaches:**

   A- **Structuralism:** Language is a static phenomenon. It does not change. (01pt)

   B- **Generativism:** Language is a dynamic phenomenon. It is generated by finite abstract rules (linguistic competence). (01pt)

4. **Sociolinguistic approach:**

   A- Language is a flexible phenomenon. It is subject to variation either cultural or social. (01pt)

   B- Language is affected by the environment which imposes the rules of speaking. (01pt)
Such an answer has been used as a model in scoring third year students exam papers, but without the addition of points that have been allotted just later. The analytic scoring method which had been adopted as a psychometric system in order to reach an objective scoring procedure had lacked two possible factors: the relevance of content and the organisation of the main ideas of discussion. As far as the relevance of the content is concerned, the possible way to control the mark given to the content is first to divide the content into main ideas and or two to give each idea a specific point. The teacher makes a general reading of the whole answer. He proceeds by reading the answer starting from the beginning to the end. The twofold objectives of such a reading are first to see whether or not the student writes on the subject, and second to look for the points of discussion that he is asked to deal with.

The student's answer can be evaluated according to four attributes as mentioned previously in the criteria of assessment. Therefore assessing the relevance of content may reveal if the essay is either knowledgeable, minimally answers the question, includes some knowledge of the subject, contains limited knowledge of the subject or misdirected and off-topic. If the answer is acceptable the teacher starts by giving each idea a specific mark according to the essay answer outline he has set forth before administering the exam question. Such an answer is used as a model. For instance, he may allocate 06/06 if the answer is recognised as knowledgeable because the student possesses a superior understanding of the question. All the ideas, as a matter of fact, have been introduced in a very precise way. This leads no doubt in the reader’s mind that the student has fully understood the question matter.

In other cases, the students may not respond to the question with the whole ideas. He may minimally answer the question or the answer may contain just some or limited knowledge of the subject. In accordance with the essay answer, the teacher may provide for instance 04/06 to an answer which lacks two ideas if the exact answer should be based on six ideas. He may
offer 03/06 to an answer which lacks three ideas and 02/06 to an answer which lacks four ideas and so one.

However, if the essay answer is formally wrong, the student will be given a mark agreed on in advance and set forth as misinterpretation of the subject. Such a mark is, according to the assessing scheme 01/20. The student's exam paper is automatically dismissed. The response is completely off-topic whatever its writing quality. The reader is left with certainty that the student has not understood the question. In such a situation, the microstructure is not to be applied. The analytic assessing procedure starts with the content and ends with it whenever the message is deemed unacceptable.

4.4. Microstrutural scoring procedure:

The second step in assessing the essay answer deals with the six other criteria which are chronologically organised according to the assessing procedure. The Macrostructure assessment procedure opens the door for the second step of the process. The student's answer seems acceptable according to the teacher's knowledge of the subject. The Microstructure assessing practice proceeds by dividing the essay answer into six assessing criteria. Each of them is to be assessed separately. We are, thus, confronted with six procedures which can be dealt with as follow:

4.4.1. Assessing the paragraphing of ideas

According to the teacher’s knowledge of the subject, the answer should follow a clear and acceptable organisation of ideas into acceptable paragraphs. The teacher forecasts in advance all the possibilities that may occur in organising the ideas of the essay answer. He tries to concentrate only on the organisation of ideas of the relevance of content into paragraphs without any reference to the other criteria such as, coherence, diction, grammar, mechanics and style. The main objectives in doing so is first to ensure all the possibilities in organising the ideas that the students are likely to adopt, and second to limit the assessing procedure to
just one aspect. The teacher knows in advance all the possibilities that can be used. For instance, the essay answer, as it is demonstrated above, consists of six ideas. These ideas can be organised into different paragraphs. There are, however, five possible ways that can be formally used as it may be illustrated in the following table. The table below displays the students’ first possible organisation of their essays’ content as expected by the teacher.

<table>
<thead>
<tr>
<th>The content expected organisation (1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>The first possible way: <strong>Two paragraphs</strong></td>
</tr>
</tbody>
</table>

A. The first idea (1) with the second (2) in the first paragraph.
- The third idea (3 / A+B) with the fourth (4 / A+B) in the second paragraph.

B. The first idea (1) the third (3 / A+B) in the first paragraph.
- The second idea (2) with the fourth (4 / A+B) in the second paragraph.

**Note:** It is possible to reverse the order. The second paragraph becomes the first in both A and B.

Table 4.1. The teacher expected content organisation (1)

The following table displays the students’ second possible organisation of their essays’ content as expected by the teacher.

<table>
<thead>
<tr>
<th>The content expected organisation (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>The second possible way: <strong>Three paragraphs</strong></td>
</tr>
</tbody>
</table>

A. The first idea (1) with the second (2) in the first paragraph.
- The third idea (3 / A+B) in the second paragraph.
- The fourth idea (4 / A+B) in the third paragraph.

B. The fourth idea (4 / A+B) in the first paragraph.
- The third idea (3 / A+B) in the second paragraph.
- The first idea (1) with the second (2) in the third paragraph.

**Note:** The fourth idea (4 / A+B) can be in the second paragraph.
- The third idea (3 / A+B) can be in the third paragraph.
- The first idea (1) and the second can be either in the second or in the third paragraph.

Table 4.2. The teacher expected content organisation (2)
The next table displays the students’ third possible organisation of their essays’ content as expected by the teacher.

<table>
<thead>
<tr>
<th>The content expected organisation (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>The third possible way: <em>Four paragraphs</em></td>
</tr>
</tbody>
</table>

<p>| | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A.</strong></td>
<td>The first idea (1) in the first paragraph</td>
<td></td>
<td></td>
<td></td>
<td>The third idea (3 /+B) in the second paragraph</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-The second idea (2) in the third paragraph</td>
<td>-The fourth idea (4 /A+B) in the fourth paragraph</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>B.</strong></td>
<td>The third idea (3 /A+B) in the first paragraph</td>
<td>The first (1) idea in the second paragraph</td>
<td></td>
<td>The fourth idea (4 /A+B) in the third paragraph</td>
<td>The second idea (2) in the fourth paragraph</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>C.</strong></td>
<td>The second idea (2) in the first paragraph</td>
<td></td>
<td>The fourth idea (4 /A+B) in the second paragraph</td>
<td>-The first idea (1) in the third paragraph</td>
<td>-The third idea (3 /A+B) in the fourth paragraph</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>D.</strong></td>
<td>The fourth idea (4 /A+B) in the first paragraph</td>
<td>-The second idea (2) in the second paragraph</td>
<td>-The third idea (3 /A+B) in the third paragraph</td>
<td>-The first idea (1) in the fourth paragraph</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.3. The teacher expected content organisation (3)

The table coming table displays the students’ fourth possible organisation of their essays’ content as expected by the teacher.
The content expected organisation (4)

The fourth possible way: *Five paragraphs*

- The first idea (1) with the second in the first paragraph.
- The third idea (3 /A) in the second paragraph.
- The third idea (3 /B) in the third paragraph.
- The fourth idea (4 /A) in the fourth paragraph.
- The fourth idea (4 /B) in the fifth paragraph.

**Note:** it is possible to reverse the order: The fourth idea (4 /B) can be in the first paragraph, the fourth (4 /B) can be in the second…etc.

**Table 4.4. The teacher expected content organisation (4)**

The following table exhibits the students’ fifth possible organisation of their essays’ content as expected by the teacher.

The content expected organisation (5)

The fifth possible way: *Six paragraphs*

- The first idea (1) in the first paragraph.
- The second idea (2) in the second paragraph.
- The third idea (3 /A) in the third paragraph.
- The third idea (3 /B) in the fourth paragraph.
- The fourth idea (4 /A) in the fifth paragraph.
- The fourth idea (4 /B) in the sixth paragraph.

**Note:** It is possible to reverse the order: The fourth idea (4 /B) can be in the first paragraph …and the first idea can be in the sixth paragraph.

**Table 4.5. The teacher expected content organisation (5)**

The organisation of ideas is an important factor in responding to an essay question. It denotes the student’s ability to put facts into a sequence of reasonably ordered paragraphs.

The procedure in assessing this category consists of the following steps: First, the teacher tries to count the number of paragraphs written by the students. As it is demonstrated above
the number of paragraphs should signal the number of ideas. There are six ideas which can be built up into two, three, four, five or six paragraphs, according to the matter of discussion.

Second, the teacher compares the ideas expressed by the students with the number of paragraphs. The student will be given a full mark, i.e., 02/02 if the number of paragraphs signals the ideas of discussion. The essay should have a discernible beginning, middle and end. The student will be given 01/02 if the essay is somewhat loosely paragraphed. The paper has minor weaknesses in paragraphing, but it contains evidence of the writer’s ability to organise information into fluent and unified paragraphs. The paragraphing is limited in logical development of ideas into paragraphs. There is, however, a serious trouble in introduction and conclusion. The student does not know how to introduce and end the ideas that logically focus the reader’s attention into the thesis or interestingly complete the development of the thesis and build up to the main points of the essay.

Third, the student will be given 0, 5/02 if there is serious difficulty in managing the tasks of the assignments. The essay seems also to lack an overall plan with a beginning, middle and end. The paragraphs are somewhat disorganised, and difficult to decipher.

**4.4.2. Assessing the organisation and coherence of ideas:**

Following the ideas of discussion and their organisation into unified paragraphs, the students are also judged for their ways to come to communicate such ideas in a more comprehensible and accurate writing framework. The assessing procedure deals with the three following attributes: First, the student will be given 02/02 if the ideas that he introduces are cogently developed. They seem logically organised and connected with clear transition. There is also a good control of language to convey ideas with reasonable quality. Second, the student will be given 01/02 if the ideas that he develops into paragraphs are well organised, but they are not connected with transitions. Third, the student will be automatically penalised with 0, 5/02 if the essay seems chaotic confused and mis-handled with no ideas connection.
4.4.3. Assessing language form

Language form consists of the three interrelated components: diction, grammar and mechanics. The assessing procedure is done at the same time. The three components are judged all together. The classroom teacher starts first by indicating the wrong words by symbols. Each symbol specifies the type of error. For instance, the errors made in diction are specified by W.C (i.e., word choice), the mechanical errors are specified by P. C. S (i.e., punctuation, capitalization, spelling) and grammatical errors are indicated by AGR, VT, X, A.W, W.O, W.F, V.F, ART, PREP, REF, (i.e., subject and verb agreement, verb tense, unwanted word, word missing, word order, word form (noun-verb-adverb-adjective), verb form (gerund-participle), article misused, preposition misused and reference (unclear relationship between pronoun and its antecedent).

In the second step the teacher starts by counting errors. Marks are reduced or deducted from the general marks allotted to diction, grammar and mechanics. For instance, in grammar the student will be given a full mark, i.e., 02/02 if the essay demonstrates consistent mastery of language. It is virtually free from errors of grammar, and there is evidence of superior control of language. The student is to be penalised for each error he will make. The subtraction is to be applied whenever the student makes an error. For each error 0, 5 is deducted from the general mark allotted to grammar. The student will be given 00/02 if he makes four errors. The same procedure is used with diction and mechanics.

4.4.4. Assessing the style

The teacher leaves the assessment of style at the end of the procedure. He makes use of the results obtained in the assessing of language form in order to judge the quality of the answer. The assessing procedure can be regarded from the following steps: First, the student will be given 03/03 if the essay is clear and free of errors. This means that the essay demonstrates mastery of elements of effective writing. It demonstrates an effective range of sentence
clarity, variety of word-choice. Second, the student will be given 02/03 if the essay is comprehensible, but some occasional errors in grammar, mechanics and diction affect the meaning of the sentence. For instance the student will be given such a mark in style if he makes less than four errors in grammar, mechanics and diction. Third, the student will be given 01/03 if the essay has the same attributes as in model (2) in addition to some ambiguous words and sentences. The essay demonstrates virtually no sentence variety. The meaning of the sentence is ambiguous due to the inappropriate or non-idiomatic word-choice. Fourth, the student will be given 0, 5/03 if the essay is chaotic due to some frequent errors and ambiguity of some words and sentences. The writer’s control of language may be imprecise, awkward, or clumsy. For instance, such a mark will be given to a style which contains more than four errors in grammar, mechanics and diction, in addition to the existence of some words and sentences which are incomprehensible and even unreadable.

The assessing procedure permits the teacher to judge the quality of each s component by a mark. At the end of each written test, he counts all the obtained marks and mentions them in an assessing grid. For instance, a student may be given 07/20. The scoring grid may be used to permit the student to know exactly why he has been given such a mark thus he may be aware of his lacunae as it is displayed in the following table. The table exposes a grid to be used in assessing every single written essay.

<table>
<thead>
<tr>
<th>The assessing grid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criteria</td>
</tr>
<tr>
<td>1. The relevance</td>
</tr>
<tr>
<td>2. Paragraphing</td>
</tr>
<tr>
<td>3. The organisation</td>
</tr>
<tr>
<td>4. Diction</td>
</tr>
<tr>
<td>5. Grammar</td>
</tr>
<tr>
<td>6. Mechanics</td>
</tr>
<tr>
<td>7. Style</td>
</tr>
<tr>
<td><strong>Total</strong></td>
</tr>
</tbody>
</table>

Table 4.6. The teacher’s assessing grid example
Such a mark (07/20) signals the weakness of the essay. It may have the following attributes: First, the essay demonstrates some knowledge of the subject. It almost answers most part of the question. It the question needs six ideas of discussion; the student has demonstrated only three of them.

Second, the essay lacks an overall plan with the beginning, middle and end. There is serious difficulty in managing the tasks of the assignments. The paragraphs are somewhat disorganised. The key ideas in paragraphs lack development or illustration. They provide little evidence of the student’s ability to develop an organise response.

Third, the essay is chaotic, confused and mishandled. There are no connections between ideas, which to some difficulties to understand the answer.

Fourth, the essay contains some errors in diction. There is an inappropriate or non-idiomatic word-choice. Fifth, the essay suggests lack of competence. It has serious and frequent problems in the use of language and sentence structure. It contains numerous errors in grammar, usage and mechanics that interfere with meaning. Such an essay suggests that the student is unable to deal competently with the question.

Finally, the meaning is comprehensible, but there are some sentences which are ambiguous and contain some occasional errors in diction; grammar and mechanics. The lack of language competence does not affect the essay but it downgrades its quality.

4.5. The analytic assessing limitations

The analytic assessing procedure could be very effective in the sense that it assists the teachers identify in a high accurate manner their students’ deficits in many writing areas, as it facilitates things to design the appropriate and effective feedback to set up and boost the learning of writing.

Yet, the analytic assessing process is not without limitations. We may point out to three major drawbacks. First, it is blamed to be time and effort consuming. An analytic approach to
assess writing would certainly require the teacher to read and evaluate each single written paper at least seven times. He would read the paper to check up seven different components i.e. he is to read a first time to assess the relevance of the content, a second time to review the paragraphing, a third time for considering the organisation, a fourth time to gauge diction, a fifth time for evaluate the grammar, a sixth time to weigh up the mechanics and a seventh time to estimate the style.

Second, and because of the high number of the EFL students in our universities, such a procedure would never attain high rates of effectiveness, objectivity, consistency and most importantly reliability. The assessment of the very first written papers would achieve a high degree of reliability however it would never be the case when the number of papers exceeds one hundred and more. The last disadvantage to consider is that such process would not be practical. The huge number of written essays to assess would never allow the teacher to determine accurately his learners’ common deficits and hindrances. In other way, he would never have an obvious insight about the general lacunae of his learners, thus he would be unable to design a collective effective feedback as remedial work.

Consequently and in this respect, the researcher would point out the possibility to overcome all those deficiencies mentioned above through transforming the assessing scheme he suggested into a computer programme. He would also signal the existence of such kind of assessment since a long time ago. Moreover, he would present an overview of the computer assessing applications available in the market of pedagogy. However and before going any further its worth to shed a slight light on the issue of automatic essay assessment.

4.6. Automatic essay assessment

Automatic essay assessment is that the computer technology that assess and evaluate the written production. It is computer software that is designed to aid the writing teachers in the process of assessing their students’ essays. Such a technology is principally used to defeat
The use of the computer to assess the written language is not a recent issue. It is dated back to 1960s. However, it has not become a reality until the unprecedented innovation of the computer sciences in the 1980s. A group of researchers working for specific firms put forward computer programmes to assess the written language. They have asserted that their computer applications agree with human assessors to really great extents. Accordingly, computer assessing has been put into practice in a lot of important testing courses. In the United States of America, for example, the use of the computer was firstly introduced to assess the writing components of the Graduate Management Admissions Test (GMAT). Such a test, mainly used in the admission to a graduate management program, is now assessed by a teacher and by software rather than by two teachers and shows very elevated rate of reliability and consistency.

Consequently, the researcher wishes to draw the attention of the EFL teachers in general and the EFL writing teachers in particular toward the existence of such computer programmes able to assess and determine in a mathematically precise way the students’ lacunae and the extents to which they are improving their writing skills. Thus, things will be simplified as far as designing an effective feedback will be concerned. In other words, automatic assessment helps the teacher design an effective feedback for a great number of students. In this respect, Ruth Breese (2012) determines the aims of the feedback saying:

Although feedback is traditionally associated with lavish use of the red pen, it is important to remember that the main purpose is
to provide a channel for teachers to communicate constructively with students and help them to develop as writers.

**Ruth Breeze (2012:139)**

It is obvious enough that the feedback is kind of facilitator or a vehicle through which the teacher may set up the whole process of teaching. Hattie and Timperley argued also that an effective feedback can have a major effect on both learning and achievement (2007:57). They postulated that the ideal feedback should include three responses: Feed up, feedback, and feed forward, that is to say, feedback should provide the student with information concerning, the learning goals, student performance, and ways of improvement. Feedback is not only for learners to bridge the breach between the actual and the wanted skills. It is also vital for educators to assess, regulate, and augment their teaching exercise efficiency and success.

The purpose of this second part of recommendations is to provide an overview of existing methods of automated writing assessment. The researcher tries to portray the most broadly used writing assessment computer systems namely the Project Essay Grader, the Intelligent Essay Assessor, E-rater, IntelliMetric, and Bayesian Essay Test Scoring System.

### 4.6.1. Project Essay Grader (PEG)

Project Essay Grade was designed by Ellis Page in 1966 to achieve more practicality, reliability and effectiveness for the essay assessing process involving a large number of students. PEG makes use of association to foresee the intrinsic value of the essays. Page utilizes the item ‘trins’ and ‘proxes’ to illustrate the way PEG evaluating and assessing a written piece. ‘Trins’ refer to the basic components such as grammar, diction, fluency, and punctuation. On the other hand, ‘proxes’ designates the connection of the intrinsic variables. Thus, ‘proxes’ refer the appropriateness of the use grammar rules, vocabulary, etc.

Project Essay Grader is very effective in allocating marks that are similar to those of human assessor. Besides, the application can computationally determine and identify the writing inaccuracies committed by the learners. Nevertheless, PEG was condemned for
paying no attention to the semantic facet of the written language and spotlighting more the exterior or the surface structure (Chung & O’Neil, 1997). Simply put such an assessing machine fails to identify the content related aspects of a written piece or an essay i.e. its relevance, organization, style paragraphing and etc. The Project Essay Grader does not supply writing feedback to learners. It was criticized to be weak as far as rating exactness as concerned. It was possible to trick the programme through developing very long essays since it was based upon an indirect assessment of the writing skill.

4.6.2. Intelligent Essay Assessor

Intelligent Essay Assessor, IEA examines and assesses writing by means of a semantic essay-analysis method. Such a method originated the ‘Latent Semantic Analysis’ approach put forward by the following psychologists Thomas Landauer, Peter Foltz and Darrell Laham.

Foltz sees ‘Latent Semantic Analysis’, LSA as a statistical model of word usage that permits comparisons of the semantic similarity between pieces of textual information (Foltz, 1996:2). The function of Intelligent Essay Assessor involves is very simple. In order to weigh up the general quality of an essay, it necessitates to be trained on well written texts. Then, the written essay needs to be typified through vectors as mathematical representation of the essay. Finally, the theoretical relevance of the content of the targeted essay is compared to other texts (Landauer et al., 1998).

The Intelligent Essay Assessor considers much the content related characteristics rather than the form related ones. But it does not mean that the Intelligent Essay Assessor disregards formal aspects such as grammar and mechanics. In other words, even though the IEA applies an LSA approach to assess chiefly the quality of the content of an essay, it also provides feedback on grammar, style and mechanics as it is clearly noticeable in the following feedback figure generated by the so-called Intelligent Essay Assessor. The following diagram displays an assessment conclusion generated by the IEA.
4.6.3. E-rater

The E-rater or the electronic essay rater was meant to determine the linguistic components in a written text. It was established by the Educational Testing Service (Educational Testing Service is an educational testing and assessment organization in USA. It has put forward many famous tests such as TOFL). E-rater exploits the natural-language processing method. The E-rater was conceived to spot definite lexical and syntactical hints and prompts in a written paper (Burstein, 2003 and Kukich, 2000).

E-rater employs a corpus-based approach. E-rater is trained on a corpus of written works assessed previously by at least two teachers on a 6-mark scale to generate a sample, in which actual essay data are used to examine sample essays. The architecture of the E-rater encompasses three elements mainly the syntactic, the discourse and the topic-analysis. Those three elements generate the final assessment of the essay. Each component considers a specific area. Consequently, the syntactic element considers the use of the syntactic structures such the use of the clauses. As for the discourse component deals with linking words and conjunctions to identify how well the essay is organised. Finally, the topical analysis element spots lexis usage and topical content.
4.6.4. IntelliMetric

IntelliMetric makes use of three main computer technologies: artificial intelligence (AI), natural language processing (NLP), and statistical technologies. IntelliMetric is conceived to adopt the human wisdom. In other words, it is developed in a way to comprehend natural language thus being able to assess any written language. Consequently, IntelliMetric deals with the essay according to the key features of standard written English. IntelliMetric is made able to store each rate related to specific characteristics in an essay answer. It is also claimed that the scoring system “learns” the characteristics that human raters likely to value. However, IntelliMetric needs to be trained with a set of pre-assessed written essays with marks allocated by writing teachers. These samples are used as a basis to deduce the assessing scale and the human assessing attitudes as well.

IntelliMetric is made able to assess about 300 semantic, syntactic, and discourse components in a written piece through the use of Artificial Intelligence and Natural Language Processing technologies. To put it simple, IntelliMetric is conceived to simulate the human brain. Thus it may imitate the human method in assessing the essay. It is based upon a difficult scheme of information processing. IntelliMetric works in many dimensions. It is recursive. It uses different assessments based on binary system. Moreover, such an assessing tool may assess essays written in many languages such Dutch, French, German, Italian, Arabic, Japanese, Spanish, and of course English.

4.6.5. Bayesian Essay Test Scoring System

The last automatic assessing programme to be tackled is labelled the Bayesian Essay Test Scoring System suggested by Lawrence M. Rudner. BETSY is somehow different from the previous tools since it may be recognized as both an assessing and a research tool together.

Such a system is based on the Bayesian theorem approach. Bayesian approach have manufactured many programme such as recognizing spam and other undesired e-mails
based on their similarity with previously classified e-mail. Bayesian is available under two samples used in assessing texts: the Multivariate Bernoulli sample and the Multinomial sample. The former assesses the existence of specific components in a written paper. The latter verifies the various use of a definite component in an essay. The Multinomial form works out quite quickly compared to the Bernoulli form.

The Bayesian approach uses key concepts such as stop words, feature selection, and stemming. Stop words point out prepositions, pronouns, adjectives, and various articles. Stemming represents the procedure of removing suffixes to find stems. For example, obtaining “form” as a stem for formation, transformation, transformational, formative and formed. Feature selection, however, indicates reducing of randomness and disorder. BETSY requires a training of 1000 written samples to grab how to assess new papers.

During the past four decades, numerous investigations took place wishing to highlight methods to combine and apply the computer technology to the writing assessment. More recently, progressively more innovative micro processor technology has allowed writing skills to be assessed automatically. In this respect, we can mention the highly developed researches of Rudner & Gagne, 2001; Rudner & Liang, 2002; Hamp-Lyons, 2001; and Attali & Burstein, 2003. Such systems of assessments do not usually assess straightly the basic components of a written essay as teachers do, but they rather employ correlations of the basic components to envisage the rate of an essay. The automatic assessing essay systems cited above employ a wide range of procedures to design an instant comment, feedback, and mark. Finally, each automatic assessing architecture requests different amounts of written essays to train the system.

4.7. Designing an effective feedback

Since the provision of feedback in EFL classrooms has always been a crucial element for both the teacher and the learner, it is not surprising that a great deal of researches have tackled
this issue. Checking students’ written works is a corner stone in almost all educational institutions where English is taught as either a second or foreign language. The writing feedback is an indispensable aspect to be considered by both teachers and learners. Throughout such a feedback, the students would recognize their concrete level of proficiency in each area of writing. Thus, they may be able to treat effectively any lacuna or deficit. On the other hand, teachers would identify in an exact way their students' weaknesses. Therefore, they may design a road map to get out of such hindrances. They would try to adjust their practices.

An effective feedback may be related to the learning goals. It may portray the strengths and weaknesses of the learners’ writing skills. And most importantly, it should feed forward. Feedback is incontestably defined as the information provided about students' performance. It is to check up the extents to which the learning goals are achieved. The feedback is meant to adjust the whole process of teaching, learning, testing, and assessing. The feedback will never be effective if it is combined neither with the feeding up nor with the feeding forward. Learners will never be able to grasp and to proceed upon the feedback they receive. The feedback should start with clear objectives that trace clearly the road towards enhancing performance. The feedback is the roadmap that tells the learners where they are coming from and what their actual location, and where they should go towards. It identifies, determines and limits in a mathematically precise way the targeted performances or competences. Setting clear goals to students will give them certain guidance about their performance.

Indeed, teachers spend much of time and efforts to read, re-read, interpret, mark and respond to students' papers. However, it seems that the feedback the teachers provide is far from being adequate and effective. The writing feedback usually takes the form of a simple and immediate teacher’s impressions about his student’s performance without any effective comments to identify the lacunae or to provide the remedy. The teachers’ impressions and
responses to the written performance are frequently negative, i.e. they often use expressions such as ‘Good’, ‘Bad’, ‘not enough’, ‘you can do better’ without neither highlighting the strengths or the weaknesses of students’ writing nor providing any remedial tasks. Besides, it seems that most of writing teachers’ feedback is focused on language errors. A few responses are provided on the content and the organization of whole writing. More significantly, feeding up and feeding forward have never made a kind pedagogical concerns for such teachers.

On the other hand, it is commonly acknowledged that students may have positive attitudes towards teachers’ written feedback. They may appreciate, consider, and understand the provided feedback. However, there is often a gap between the kind of supplied feedback and what the students would like to receive. Students usually desire both of positive and negative feedback. They prefer detailed feedback that offers a solution and assistance towards achieving a better performance in writing. Such features are often missed in the feedback presented by the writing teachers. Also, teachers do not always consider their students’ needs and suggestions. The feedback is usually shaped by time constraints and class size. The writing teachers are aware of the significance of an effective feedback; however, it seems that their thoughts and beliefs are not reflected in their actual practice.

Consequently, the researcher wish to provide few suggestions to reshape and enhance the effectiveness of writing feedback through exploring the different roles of the writing teacher when delivering such feedback to his students’ writing. The aim here is to help teachers and learners to develop a conceptual understanding of feedback issues, and thereby, enhance their learners writing.

4.8. Reshaping the writing feedback

To lessen the fissure between the learners’ perception and the expected purposes, the writing teacher would provide an effective feedback. However and before designing such a feedback, the teacher is supposed to raise three key questions:
1- Where am I going?
2- How am I going?
3- Where to next?  \textbf{(Hattie & Timperley, 2007: 86)}

These questions are related to three concepts: feed up, feedback, and feed forward. Hattie & Timperley have suggested a feedback model, illustrated in the figure below, to point out the necessity to address these three questions for an ideal learning environment or experience. \textbf{Hattie \& Timperley, 2007:88} . The figure displays the model suggested by Hattie & Timperley to design and effective feedback.

![Feedback Model](image)

\textbf{Figure 4.4. Hattie \& Timperley model of an effective feedback}

It is clear from the model that the initial point is to set up the objective. Then, to ask the three mentioned above questions: “Where am I going?” to refer to the objectives or to the feeding up, “How am I going?” to refer to the feedback, and finally “Where to next” referring
to the feeding forward. Each feedback question works at four different levels namely task, process, self-regulation and self levels.

4.8.1. Where am I going?

As it is noted before, feedback is useful information provided for both the students and teachers that is related to a specific goal. However, and before going any further, it is worth that the learners have to understand the desired end of a prearranged lesson.

Goals are in two dimensions: challenge and commitment: challenging goals inform learners as to what type or level of performance is to be attained, so that they can direct and evaluate their actions and efforts accordingly. Feedback allows them to set reasonable goals and to track their performance in relation to their goals so that the adjustment in effort, direction, and even strategy can be made as needed. (Hattie & Timperley, 2007:88). To put it simple; it informs learners about their level of performance. It also allows them to set appropriate challenging objectives to move learning forward. Objectives need to be well defined and appropriate so that an adjustment can occur. However, if the learning goals are weekly set up the feedback will never be considered as effective. The best example is one of Hattie and Timperley, students are given feedback on presentation, spelling, and quantity in writing when the stated goal is “creating mood in a story”. (Hattie & Timperley, 2007:89). Teachers and students are committed to these challenging goals; it leads to a better understanding and greater efforts and engagement.

4.8.2. How am I going?

Answer to this question provides learners and teachers with information related to a specific task. It is not enough to make learners aware about their performance but, also how to boost such a performance. The dilemma is that teachers usually confront barriers when addressing this question since students do not always like the answer. According to Hattie & Timperley, a test is one method to get a reply to this question.
4.8.3. Where to next?

The notion of feed forward answers this question. It is the use of information so as to recognize where gaps are in students' learning and use this information to move learning forward. Hattie and Timperley stated that these may include enhanced challenge, more self-regulation over the learning process, greater fluency and automaticity, more strategies and processes to work processes to work on the task, deeper understanding, and more information about what is and what is not understood. (Hattie & Timperley, 2007:90). It has been said that this feed forward question has the most powerful effect on learning and progress in general. All these three main questions have to be integrated in order to ascertain the power of feedback. Aiming to close the gap between the current state of learners and where they are aiming to be.

4.9. Enhancing the effectiveness of the feedback.

Written feedback is a pedagogical instrument to enhance learners' writing and learning in general. That’s why providing an effective feedback that learners can digest and act on is necessary. Based on the findings of research, various pedagogical implications can be presented. The main assumption here is that not only teachers have to work on their feedback practices, but also, students have to know what to do and how to use this feedback. Burke and Piterick argue that two things should be made to enhance the effectiveness of teachers' written feedback: First, tutors need to develop their awareness of issues surrounding the provision of feedback so that the feedback provided is more appropriate for student learning. The second step requires tutors to prepare their students for feedback and to help them develop strategies to unpack and act on tutor feedback. (Burke and Pieterick, 2010:06). Therefore, to make the provision and reception processes of the teacher's written feedback more effective, the following suggestions would be very useful.
4.9.1. Reflecting upon the feedback process

Teachers should have a deeper understanding about their feedback practices. In this sense, teachers need to be reflective. They need not to take the process of giving a feedback for granted. They have to examine, assess, and compare their own feedback practices with other theories. When teachers reflect on their written feedback, and monitor their practices, they become aware of what is needed to be changed and improved. And this will assist them to consider their learners and their attitudes concerning the provision and the reception of written feedback. Different procedures can be used such as asking other teachers to express their views concerning the feedback provided.

4.9.2. Considering students’ needs and suggestions

Most of the time the designed kind of the teachers’ writing feedback is always affected by time and syllabus constraints. In other words, the writing teacher’s feedback has never reached 100% of the predicted objectives. Moreover, such a feedback has never been designed according to the learners’ needs and suggestions thus it has never been reliable and effective. The sort of the writing teacher feedback often considers time restrictions, but not EFL learners’ needs and preferences.

Teachers may refer to their students' requirements and inclinations. Writing teachers could discuss with their learners about their different views concerning the quality and quantity of provided feedback. As they might also enquire about the kind of feedback their students would like to receive. Haines (2004) state that “listening to student views on feedback can guide us as to how students are likely to receive and interpret our comments” (as cited in Burke and Pieterick. 2010:72). Learner is the central point of any teaching process so he is to be involved to make any kind of decision concerning what happen inside the classroom. Involving learners in designing any teaching item is so useful, as it may motivate the learner...
and facilitate many things for the teacher. Such a teacher could tackle exactly his students’ immediate hindrances.

4.9.3. Varying feedback brands

It is undeniable that an effective feedback should include as much details as the students’ lacunae and deficits. To put it simple, an effective feedback should encompass the following four traits:

- Detailed
- Feeding up
- Feeding forward
- Balancing between positive and negative feedback

Yet, generating such sort of a writing feedback is a challenging and time consuming task. However, it is always worth to attempt doing so. Teachers may first consider the time allocated to the feedback. Using the time effectively is of a paramount significance in enhancing both teaching and feeding back writing. The evidence shows that teachers should teach less and give feedback more because feedback is a tool to make learning happen.

Moreover, it should be acknowledged that the teacher is not the exclusive writing feedback provider. Despite the fact that the learners usually share a kind of sympathy and preference towards their teacher's feedback, other feedback resources could be used to boost up the process such as self-assessment feedback, peer feedback, and electronic feedback. The teacher is to motivate and raise awareness about the usefulness and effectiveness altering the feedback source, thus giving extra dimensions to learning writing. The student-teacher conferences either one-to-one or in-groups is so significant to provide and receive the appropriate feedback. It is argued that teachers have to use different methods to cover the needs and preferences of diverse learners. Each learner may have his own specific desires, requests, preferences, aptitudes and suggestions concerning the feedback they would like to receive.
4.9.4. Feeding up

Feedback is useful information about students’ performance. And students’ performance is the extent to which learning goals are achieved. Setting explicit goals will incite students to break the gap between the actual performance and the intended one. Burke and Pieterick postulated that

The starting point for tutor feedback, then, is the learning outcome. Aligning feedback with learning outcomes means giving feedback in reference to specific types of goals so that students know if they are on the right track and to what extent they have reached their goals through the feedback the tutor gives them.

(Burke and Pieterick, 2010:34)

Feed up is not only communicated between the teacher and the learners but rather, an essential component of the written feedback. Feed up is an important part of the feedback process as it enables learners to know the relevance of teacher’s feedback and how these information can help them to close the gap between what had been done and what normally should be done. Feeding up is likely to trace a map for the learner about the road to track in order to achieve a better writing performance.

4.9.5. Equilibrating positive and negative feedback

Teachers usually focus their written comments on only negative aspects in students' writing. Others usually concentrate on only the positive aspects. In the same respect, some students prefer negative feedback as it reveals their lacunae, while others prefer the positive one as it encourages and motivates them. Indeed, both of positive and negative feedback are important for learners to improve their writing.

Teachers may mitigate their feedback through integrating the positive feedback. They may set a kind of balance between both of positive and negative feedback. In other words, teachers might soften the voice through which they deliver feedback. For example, instead of
commenting on telling a student’s spelling mistakes saying “a lot of spelling errors in your essay”, it is advisable to telling him “you have written a good introduction and conclusion, but you have some spelling errors in your essay”.

4.9.6. Feeding forward

Feed forward is the most important aspect in the process of feedback. What significance the writing feedback could have if it does not point to ways of improvement. Feedback should answer the question: “where to next”, that is to say, what the learners need to improve his writing skills in the next assignment. It should give some remedies fill the breach between the actual performance and the desired one. For instance, the writing teacher could point out the importance of generating short sentences to enhance the overall structure of the essay. Moreover, he could advise his learners to read intensively to enrich their vocabulary. He could also invite them to use their electronic devices to check up spelling of complicated items thus avoiding such kind of mistakes and errors.

On the other hand, the teacher is to define and specify his writing feedback. The writing feedback should be personalized since it is a response to a person, not to a text, that s’ why teachers should address the feedback by using names and pronouns such as “you”. For instance, instead of saying “A lot of spelling errors”, the teacher could say “You have made a lot of spelling errors”. Moreover, he should point to a specific area that needs an improvement. Feedback like “Details are missing” is to be avoided unless it includes information on which and how the details could have been included in an effective way.

4.9.7. Incorporating remedial task

Teachers can integrate different activities to their writing sessions so as to make students digest and act on the feedback provided to them. One of the activities is to provide the student with their feedback comments and ask them to read it and write two sentences on what they will change in the next assignment. They can also re-write their essays using the given
feedback and compare them. Consequently, the students would recognise how important feedback is. Besides, they would surely make a good use of their teacher’s feedback. Finally, it is advisable for the teacher after each writing test and assessment to design a remedial task allocated to correction and treatment of both common and individual errors.

4.10. Planning a corrective session

The researcher adopted the analytic assessing procedure, as displayed in the figure below, which permitted him to follow step by step the assessment of each composition.

![The assessing procedure](image)

**Figure 4.5. The assessing procedure**

The analytic assessing procedure consisted of dividing the procedure into six steps according to the number of the assessing criteria. The classroom teacher starts by judging the quality of the answer according to his knowledge of the subject-matter. The second step is mainly dependent upon the first one. The correctness of the content is to be regarded as a starting point to the other criteria. The relevance of the essay answer is determined by whether the student has come to understand the question or not. If the answer is regarded off-topic, all the other criteria are deemed irrelevant. The teacher, thus, relied on his assessing procedure on the basis of the relevance of the ideas expressed. The following criteria are also judged step by step. Each writing test was assessed according to the proposed method. The three steps of assessing were based on a psychometric basis. In each step a counting procedure was used.
Each assessing criterion was first judged and allotted a specific mark. At the end of each assessment procedure an assessing grid had been used. This assessing grid consists of the all the assessing criteria with the allotted marks as the following table may indicate.

<table>
<thead>
<tr>
<th>Students</th>
<th>Relevance</th>
<th>Paragraping</th>
<th>Coherence/Organisation</th>
<th>Diction</th>
<th>Mechanics</th>
<th>Grammar</th>
<th>Style</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>06/06</td>
<td>03/03</td>
<td>02/02</td>
<td>0,5</td>
<td>0,50</td>
<td>01</td>
<td>00</td>
<td>07</td>
</tr>
<tr>
<td>2</td>
<td>05</td>
<td>01</td>
<td>01</td>
<td>0,5</td>
<td>01</td>
<td>0,5</td>
<td>00</td>
<td>10</td>
</tr>
</tbody>
</table>

Table 4.7. The researcher’s adapted assessing grid

Consequently reading through the students’ deficits in writing would be an easier task. Skimming through the above table would certainly tell you many details about not only the whole classroom learners’ strengths and weaknesses but also about each individual inside the class. The unique trouble with such a grid is its time and effort consuming aspect.

4.11. Managing the corrective conference

A corrective conference is a compulsory practice after undertaking an assessment procedure. However, the writing teachers are rarely sparing a fraction of time to embark on such kind of a corrective session. The corrective conference lies at the heart of the whole process of teaching, testing, assessing and providing feedback to every subject particularly to writing. Assessing the students written productions without attempting a corrective conference leads no way. It is similar to someone who is very ill. He goes to the doctor. After the diagnosis, the doctor prescribes him some medicines. But unfortunately, that patient would never take such treatment. Consequently, he would never be cured. When providing his feedback, the teacher is prescribing some solutions and remedies to actual students’ writing troubles. And the corrective session seems in this respect to be the consumption of such specific remedy through drills and many various activities.
The corrective gathering is adapted by the teacher according to the conclusions he made after assessing his learners’ writing papers. Therefore, testing, assessing and scoring the writings, it is worth to review everything directly or indirectly related to the essay’s topic during the so-called corrective conference. When assessing the written works, the teacher may include a checklist to record not only all the common mistakes but also each learner’s error or what it is referred to as individual error in order to re-teach and readdress everything. The corrective session could be achieved through the following suggested steps:

4.11.1. The review of the essay question

It is a kind of a warm up to the corrective class. So, before embarking on corrective tasks or any remedial work, it is advisable to introduce and pave the way to the whole session through calling back the essay question focusing on what the learners had to deal with. During this phase, the teacher may remind his students about the written expression outline that they had to tackle and how they had to do. The teacher may even design the whole essay plan or outline on the blackboard. This stage could be achieved in different ways according to the teachers’ savoir-faire, the classroom size, the time allocated to the session, and so on.

4.11.2. Teacher’s impressions

The teacher could comment his students’ written works without a real emphasis on the errors. The comments in this phase could be broad and general. The teacher may inform his learners whether they have done a good work or not. Besides, he may stress their strengths as well as their weaknesses inciting them to make extra efforts in specific areas. He may, even, point to a good written work and thank and praise its writer to motivate and to create a sense of competition among students.

4.11.3. Collective Correction

Collective correction announces the beginning of the remedial work. It is an extremely vital part of the process of teaching writing since it deals the student’ common deficits and errors
Collective correction has, as aim, to help the student deal with their errors to avoid them later. Teachers may proceed through collective correction in different ways but very briefly. Making mistakes would be seen as a normal fact and a sign that the learning process is taking place. The researcher suggests two ways to proceed through the collective correction stage.

First, the collective correction could be done through the collective elaboration of a new essay. The whole classroom students could have the opportunity to contribute to the writing of such a new essay. This method strongly requires the participation of all the students together. The teacher may activate the voiceless learners or the shy ones and involve them in the correction process. The aim behind collective correction is to teach the students how to build and rebuild a paragraph paying attention to coherence i.e. the semantic structure unity and to cohesion i.e. the sentences structure unity. It is a step-by-step writing since the teacher would deal with each paragraph separately mainly the introduction, the body, and the conclusion. The teacher is to go very slowly to allow a better understanding. Besides, he could mark a pause at the end of each paragraph to offer the students an opportunity to review, to check, to question or to suggest something else. It is worth mentioning, here, that the elaborated essay should be brief, short and simply written to make the students remember the essay main headings, and even memorise some expressions and structures to be reused and recycled in any future assignment.

Second, the collective correction could be done through correction of a selected classmate essay sample. This technique consists on a careful selection of one student’s written piece sample. The selection should be very qualitative. In other words, the teacher is to look for the written essay with the most common errors found in approximately every assessed paper. Moreover, the selected sample should include many sorts of errors thus having the opportunity to deal with various deficits. The teacher may either write down the whole selected essay on the blackboard or make some copies for every learner (The mark allocated
to the selected paper should be hidden). Then, the teacher reads the paragraph slowly and loudly for the class. Few students may be asked to read again the essay (if time allows). Next, the teacher is to call his learners to express their opinions and impressions about the essay to assign a first mark. The class would agree for a single mark. The teacher, then, asks the students to ensure if the essay deserves the allocated first mark or not. At this moment, the valid correction starts. The teacher would proceed at two levels: first, at the sentence level to correct grammatical and lexical items, and second, at the paragraph level to check the orderly, continuity and consistency of facts. The most important fact is that students might always be involved in each step of the process. The students are to assess and correct their friend's essay under the guidance of the teacher.

4.11.4. Individual Correction

Since we have always to go from the general towards the specific, it is advisable after treating the collective errors to shift toward the individual ones. This stage may be very imperative in the suggested corrective process. It aims at making each individual student aware and conscious of the mistakes he is making. Besides, each student is to be equipped in this phase with efficient and sufficient tools to overcome his own mistakes and to solve any trouble that would stand in his way towards a successful writing performance. And, since the teacher would have recorded each student’s errors in his checklist, he could point directly to the learner to deal with his individual weaknesses. Thus, he would improve the whole class performance. However, it might be mentioned that individual correction is too demanding and time-consuming especially with large size classroom. Therefore the writing teacher may deal only with the students with the highest errors number.

4.11.5. Remedial work

Remedy, as its name indicates, is meant to treat the students’ weaknesses by supplying some remedial exercises. Two kinds of exercises could be dealt with: those which concern the
coherence of writing, and those which concern the cohesion of writing. Remedial task is done separately during another session. The students are invited to reinvest everything they have dealt with during the corrective session. The remedial work is a kind of summative assessment.

4.12. Conclusion

The essay-type test, as it is always referred to, is subjective in the sense that it requires more personal judgements rather than objective interpretations. Such subjectivity is still a problematic which is the result of individual standards of the grader and unreliability of assessing procedures. However, in order to put an end to such a problematic some guidelines have been suggested as a tentative solution to the formal assessing of an essay test question. The assumption that assessment unreliability is caused by faulty assessing procedure has lead the classroom teacher to be more aware about the extraneous factors which are closely related to the measurement setting and to variations in personal characteristics rather than to the outcome being measured. To the extent that the assessing procedure is free of such extraneous influences, the results (and decision based on the results) will be consistent and stable.

Accordingly, reliability can be increased if we standardize the administration of the assessing procedure. The proposed guidelines are set forth as testing measurements for practical purposes. They are not supposed to have attained reliability in its exact sense or even a substitute for it such as probability. We constantly fall into the same error of supporting that an absolute solution to the assessment of assays is attainable and we look for revelation. In this respect the computer seems to do task better than human assessors since no subjectivity or prejudices may be demonstrated from the part of the machine. Many computer programmes have been set forth in order to overcome the reliability and effectiveness of the assessing procedure. Automatic essay assessment is a set of computing programmes developed and intended to assist teachers assessing their students’ written productions. Their key objectives
and roles can be recapitulated in attaining reliability and validity, defeating time constraints, and overcoming many controversial issues related to writing assessment.

Achieving objectivity and reliability through either the computer or the human assessment would no doubt result in an effective feedback. Once recognizing and determining his students’ common deficits, the teacher could design an effective feedback. Written feedback is an essential aspect to be considered by both of teachers and learners. It is through feedback, students recognise their level of proficiency and how to improve it. Also, teachers would identify their students' weaknesses and try to adjust their practices. An effective feedback should be related to the learning goals. It should highlight not only the students’ writing weaknesses but also their strengths. And most importantly, it should feed forward. Indeed, teachers spend much of time and efforts to read, re-read, interpret, mark and respond to students' papers. However, it seems that the feedback the teachers provide is far from being adequate or an effective one. To provide an effective written feedback on students' written work is a difficult and time-consuming task. However, it is an important part of the teaching and the learning process. The effectiveness of feedback can be achieved only if the teacher examines the process of the provision and the reception of written feedback and find remedial measurements to redress the whole situation. As a matter of fact designing a corrective conference may solve many writing related troubles through its collective and individual corrections. During a corrective session, the learners would be not only exposed to their immediate mistakes but also equipped with means to get rid of such barriers.

All in all, every technique in teaching, testing, assessing, and feeding back written responses remains tentative forever, but we shall not cease from exploration. We must admit that the end of our exploring will be to arrive where we started and know the place for the first time.
General Conclusion
General Conclusion

Nowadays, English is incontestably the globe language. It is also the language of science and technology. English has witnessed successive phases of success and growth for nearly three centuries, i.e. since the beginning of the first industrial revolution. English has become the dominant language in numerous spheres such as education, politic, trade, technology, science, communication and many others. 80% of computer data are written in English and out of 400 million Internet users, 46% communicate in through the same language. Ability to communicate in the language of Shakespeare is one of the most requested skills by the worldwide employers.

Consequently, a giant number of people are wishing to learn such hegemonic language. They are taking part in countless learning programs all over the world. Yet, dealing with such a language is not an effortless task. Researchers are working hard to enhance the situation and to make English learning more than just simple. Teaching English has become a prime challenge for teachers. They are to teach, assess, score, provide feedback then decide next steps in learning. Learning English, as demonstrated by many linguists goes through the learning of four skills namely the listening, the speaking, the reading and the writing skills.

Writing seems to be the thorniest skill to acquire since learning to write is like learning to play piano. It is too hard and too demanding. You may sing, listen to music, appreciate songs, but to play piano is too difficult. You may read, speak, understand a language perfectly, but you may not write correctly. Teaching how to write effectively is one of the most important lifelong skills teachers report to their students. Educators must be sure to select resources and support materials that not only aid them in teaching how to write, but will also be the most effective in helping their students learn to write. Writing is the primary basis upon which the work, the learning, and the intellect will be judged in college, in the workplace, and in the
community. It expresses who someone is as a person. It fosters the aptitude to explain a complex position to readers and helps teachers designing feedback. In addition writing help refine ideas. Besides, writing requires that the learner anticipates his readers’ needs. The ability to do so demonstrates the intellectual power. Furthermore, writing stimulates the writer to expand a line of thought beyond his first impression.

Learning to write is to a great extent dependent upon the effectiveness of the teachers’ feedback. In its turn this latter is completely based on the validity and reliability of the kind of assessment provided. The conventional assessment of the learners’ writing still has a justifiable lay in the English language skills, but could not be the secure means of assessing writing. Assessment is imperative for an obvious and reliable portrait of how learners are recovering and how fine the processes of instructions deal with the learners’ requests. The assessment of essays can take two shapes either the holistic or analytic.

In holistic evaluation, the written production is read for a broad judgement and according to such a judgment the teacher assesses his learner’s levels of proficiency. All the facets of the composition (content and conventions) influence the teacher’s response, but none of them is particularly recognized or frankly addressed through a checklist. Teachers can never graph students’ progress and collect information that will assist them in defining rate and worth. This approach is speedy and proficient in reviewing and checking the overall performance. It may, yet, be unsuitable for estimating how well learners have applied precise standards or developed a particular form.

The analytic assessment, however, regard writing as to be composed of a range of characteristics, such as relevance, grammar, organization of ideas, expression of concepts, and punctuation, each of which is to be assessed independently. The essays are to be read according to a prearranged list of criteria. Those essays can be measured up according to a set
of principles. The teachers can design reliable assessing system based on selected criteria of evaluation in order to assess each written production aspect alone. Accordingly, the assessment method can guarantee a positive and constructive feedback. Analytic assessing could help the teacher keep the full of writing features in mind as he assesses the written language. It also let the learners see areas in their essays that need work when accompanied with written commentary and remedy. Its diagnostic nature offers learners a road map for an effective enhancement and perfection. However, before embarking on the analytic assessment, the writing teacher is to select the assessing components. Many researchers agreed that the ability to write encompasses minimally six ingredients mainly the grammatical skill or the ability to write English in grammatically correct mode, the Lexical talent or the skills to properly select and employ the words, the mechanical cleverness or the aptitude to rightly use punctuation, spelling, capitalisation, etc., the stylistic proficiency or the faculty to use sentences and paragraphs fittingly, the organizational skill or the capability to arrange the written work according to the conventions of English, including the order and range of material, and finally the judgement of appropriateness or the ability to make judgements about what fitting depending on the task, the purpose of the writing, and the audience.

To compare both, the holistic review is characterised by its easiness. Nevertheless the analytic assessment appears impossible to attain such attribute. Besides, the analytic procedure as opposed to the holistic one is recognized as time consuming. The teachers who assess analytically are usually required to make many separate judgments about one piece of writing. The writing teacher would spare a great deal of time and effort because of the limitless criteria of evaluation he has to assess in each essay.

In this respect and as a reflection upon the limitations of the analytic assessing procedure, the researchers suggested an involvement of computing technology as an innovative and reliable means of assessment. Automatic essay assessment is the computer technology that
assesses and evaluates the written production. It is computer software that is designed to aid
the writing teachers in the process of assessing their students’ essays. Such a technology is
principally used to defeat time, reliability, and credibility matters in the assessment of the
written language. A lot of surveys attempted to weigh up the exactness and consistency of the
automatic essay assessment programme. The findings of numerous investigations pointed out
a great conformity between the computer assessment and that of the human. The results have
shown that such technology demonstrates an important agreement rate with human assessors.
Automatic assessment of the learners’ essays has become today a reality. Computing
software are planned and designed to assist teachers in assessing their students’ written works.
Their main aims and missions can be recapitulated in realizing reliability and validity,
overcoming time constraints, and overcoming many controversial matters associated with
‘Criterion’, ‘IntelliMetric’ and ‘Bayesian Essay Test Scoring System’ are few examples of
such computing programmes that are able to assess the written production as effectively as
the teacher does.

However, many studies and investigations have drawn attentions towards many
deficiencies and drawbacks of such computing software namely the automated essay
assessment. For example, such programmes necessitate an enormous corpus of the sample
text to train the system. Besides, those programmes lack the human interaction. Moreover,
such a technology is vulnerable to cheating. In spite of all the shortcomings previously
mentioned, automatic assessments still incites and provokes the curiosity of researchers and
teachers, schools and universities all over the world.

The findings of this research have revealed that learners’ writing enhancement requires
two key deeds from the part of the teacher: First, he may make the learners distinguish and
recognize not only the writing deficits, but also their nature i.e. if they concern relevance,
style, diction, grammar, mechanic, paragraphing or coherence. Second, a competent teacher might provide his learners with adequate and effective tools to defy not only the mistakes they commit, but also the mistakes which they will commit. Pedagogically speaking, the analytical method is the most effective and formal tool to assess the written language. Its weaknesses can never be barriers to preserve the trustworthiness of such a method. Experiences have shown that the analytic method can be formally used by the classroom teacher, thus the learners’ achievements in writing are effortlessly controlled. The assessment systems as well as evaluation are closely tied to a teacher’s own personal philosophy regarding the best way of assessing. Then the writing teacher is to specify the criteria of writing assessment. A distinction must be drawn between the macrostructural assessment and the microstructural assessment. In the former, the teacher is to deal with the relevance of the essay i.e. the information and data suggested by the learner as an answer to the essay question. In the latter, the teacher is to treat every single component such as the paragraphing, the organisation and coherence, the diction, the grammar, the mechanics and finally the style.

At last but by no means least, the researcher suggested in this humble study an assessing scheme to boost the reliability of the writing assessment and thus to provide the learners’ written productions with the most possible effective feedback. Moreover, he wished to transform such a scheme into a computing programme or software to overcome the limitations of the analytic assessing method. The future suggested study would certainly be on how to convert such analytic formula into computer software thus defeating all the shortcomings mainly those related to high number of the EFL learners in our classrooms today.
References
References

Books


• Harmer, J. 2012. *How to Teach Writing*. Pearson Education Limited 2004


**Articles**


• Ferris, D.R. (2004). *The “grammar correction” debate in L2 writing: Where are we,
and where do we go from here? (and what do we do in the meantime?) Journal of Second Language Writing.


- Mahilli. (1994). *Responding to Student Writing.* ETF, 32, PP. 24-27


• Raimes, A. (1981). Composition: Controlled by the teacher, free for students. ETF, 19, 3, pp.2-7


**Webliography**


Appendices
Human Language

Language makes contact with the world on three sides: The ...........side, the........... , and the............. side. The first side deals with both Speech and.......... It is the most important feature that human beings share with each other .The second side ,on the other hand ,has no common characteristics with the first one It deals with the ............. feature that can be used to communication information. In other words, human beings are able to speak with their................bodies. The last side, as A matter of fact, falls under the heading of..........It deals with .................,and.................

The faculty of human language is entirely dependent upon such possibilities .It is a wide range of abilities that human beings possess. It is seen from three levels: ............... , ................. ............. The first level falls under the heading of...............This latter is sometimes called...................... It deals with........................, as a discipline, the..........as the subject of study .The second level , on the other hand ,is ........... .It deals with .......... , and ............The former studies the..........of a language and how they are combined to make................The latte makes use of ........for the arrangement of words into .......... and ............ In other words, it deals with............ order.

The two first levels are to be regarded as the......... of a language .Their combination in social intercourse Constitute its ........... .This latter falls under the discipline of ............Its main approach is to study Language as a means of communication .Accordingly ,the meaning of any utterance may be seen from three types: ............... , ............... and ........... .The first type is well- known as the............... Meaning .It is the logical and
formal interpretation of a word. The second is mainly regarded from four types: ……………………, ……………………, ……………………, and …………………… . It is the association of lexical items with Social factors. Language in such a situation is seen as a metaphorical tool and a …………………… phenomenon. It is a device by which individuals can maintain and establish a firm relationship with each other. The third and last type is called …………………… meaning. It is the organisation of some detailed information about a subject-matter.
Appendix 2

First Semester Exam

Essay question: As opposed to Micro-linguistics, Sociolinguistics has been established as a multidisciplinary field of investigation with a distinct approach. Discuss.
Appendix 3

Second Year Written Test Answer Sample

University of Djillali Liabes
Module: Linguistics
Faculty of Letters and Arts
Academic Year: 2013-2014
Department of English
Level: Second Year

Full name:.................................Group:........Mark:......../20

First Semester Exam

Fill in the gaps with the appropriate items. (1 mark) for each right item, and (-0.5) for each spelling mistake.

Human Language

Language makes contact with the world on three sides: The linguistic side, the Kenetic, and the Semiotic side. The first side deals with both Speech and Writing. It is the most important feature that human beings share with each other. The second side, on the other hand, has no common characteristics with the first one. It deals with the Paralinguistic feature that can be used to communicate information. In other words, human beings are able to speak with their Visual bodies. The last side, as a matter of fact, falls under the heading of Semiotics. It deals with Sign and Symbols.

The faculty of human language is entirely dependent upon such possibilities. It is a wide range of abilities that human beings possess. It is seen from three levels: Pronunciation, Grammar, Meaning. The first level falls under the heading of Phonology. This latter is sometimes called General Phonetics. It deals with Phonetics as a discipline, the Phoneme, as the subject of study. The second level, on the other hand, is Grammar. It deals with Morphology and Syntax. The former studies the Morphemes of a language and how they are combined to make Words. The latter makes use of Rules for the arrangement of words into Phrases and Sentences. In other words, it deals with Word order.

The two first levels are to be regarded as the Forms of a language. Their combination in social intercourse constitute its Meaning. This latter falls under the discipline of Semantics. Its main approach is to study Language as a means of communication. Accordingly, the meaning of any utterance may be seen from three types: Conceptual, Associative and Thematic. The first type is well-known as the Denotative Meaning. It is the logical and formal interpretation of a word. The second is mainly regarded from four types: Connotative,
Collocative, Social, Affective, and Extralinguistic/socio-cultural. It is the association of lexical items with Social factors. Language in such a situation is seen as a metaphorical tool and a Social phenomenon. It is a device by which individuals can maintain and establish a firm relationship with each other. The third and last type is called Thematic meaning. It is the organisation of some detailed information about a subject-matter.
Appendix 4

The essay answer: 06/06

The answer is based on two points:

1- Microlinguistics as an independent study versus Sociolinguistics as a multidisciplinary field of investigation

2- Sociolinguistics adopts a distinct approach which is quite different from microlinguistic approaches.

The answer must contain the following points of discussion:

1- Microlinguistics is seen as an independent discipline. It studied language without any reference to other approaches. It set up its own methods which did not share with the other disciplines. (01pt)

2- Sociolinguistics as a new field of investigation is mainly dependent upon four disciplines such as Generative Grammar, Dialectology, Anthropology, and Sociology. (01pt)

3- Microlinguistic approaches:

a- Structuralism: Language is a static phenomenon. It does not change. (01pt)

b- Generativism: Language is a dynamic phenomenon. It is generated by finite abstract rules (linguistic competence). (01pt)

4- Sociolinguistic approach:

a- Language is a flexible phenomenon. It is subject to variation either cultural or social. (01pt)

b- Language is affected by the environment which imposes the rules of speaking. (01pt)
Appendix 5

Selected Cloze Test ‘A’

University of Djillali Liabes
Faculty of Letters and Arts
Department of English

Module: Linguistics
Academic Year: 2013-2014
Level: Second Year

Full name: Benfrajd Mahdi
Group: A
Mark: 1270

First Semester Exam

Fill in the gaps with the appropriate items. (1 mark) for each right item, and (-0, 5) for each spelling mistake.

Human Language

Language makes contact with the world on three sides: The **linguistic** side, the **semantic** side, and the **pragmatic** side. The first side deals with both **speech** and **writing**. It is the most important feature that human beings share with each other. The second side, on the other hand, has no common characteristics with the first one. It deals with the **phonetic** feature that can be used to communication information. In other words, human beings are able to speak with their **bodies**. The last side, as **a matter of fact**, falls under the heading of **grammar**. It deals with **syntax** and **semantics**.

The faculty of human language is entirely dependent upon such possibilities. It is a wide range of abilities that human beings possess. It is seen from three levels: **phonology**, **syntax**, and **semantics**. The first level falls under the heading of **phonology**. This latter is sometimes called **phonetics**. It deals with **phonetics**, as a discipline, the **phonemes** as the subject of study. The second level, on the other hand, is **syntax**. It deals with **morphology** and **syntax**. The former studies the **morphology** of a language and how they are combined to make **words**. The latter makes use of **rules** for the arrangement of words into **phrases** and **sentences**. In other words, it deals with **wording** order.

The two first levels are to be regarded as the **forms** of a language. Their combination in social intercourse constitute its **meaning**. This latter falls under the discipline of **semantics**. Its main approach is to study Language as a means of communication. Accordingly, the meaning of any utterance may be seen from three types: **denotative**, **connotative**, and **contextual**. The first type is well-known as the **meaning**. It is the logical and
formal interpretation of a word. The second is mainly regarded from four types: .......œ, .......σ, .......τ, .......ν, and .......μ. It is the association of lexical items with Social factors. Language in such a situation is seen as a metaphorical tool and a .......ζ phenomenon. It is a device by which individuals can maintain and establish a firm relationship with each other. The third and last type is called .......ζ meaning. It is the organisation of some detailed information about a subject-matter.
Appendix6

Selected Cloze Test ‘B’

University of Djillali Liabes
Module: Linguistics
Faculty of Letters and Arts
Academic Year: 2013-2014
Department of English
Level: Second Year

Full name: Salmae. ................. Group: B.... Mark: 16/20

First Semester Exam

Fill in the gaps with the appropriate items. (1 mark) for each right item, and (-0, 5) for each spelling mistake.

Human Language

Language makes contact with the world on three sides: the linguistic, the social, and the personal. The first side deals with both speech and writing. It is the most important feature that human beings share with each other. The second side, on the other hand, has no common characteristics with the first one. It deals with the .......................... feature that can be used to communicate information. In other words, human beings are able to speak with their .................. bodies. The last side, as a matter of fact, falls under the heading of .............. It deals with .................................. and ..................

The faculty of human language is entirely dependent upon such possibilities. It is a wide range of abilities that human beings possess. It is seen from three levels: .................. The first level falls under the heading of .............. This latter is sometimes called .................. It deals with .................... as a discipline, the .................. as the subject of study. The second level, on the other hand, is .................. It deals with .................. and .................. The former studies the .................. of a language and how they are combined to make .................. The latter makes use of .................. for the arrangement of words into .................. and .................. In other words, it deals with ..................

The two first levels are to be regarded as the .................. of a language. Their combination in social intercourse constitute its .................. This latter falls under the discipline of .................. Its main approach is to study Language as a means of communication. Accordingly, the meaning of any utterance may be seen from three types: .................. and .................. The first type is well-known as the .................. Meaning. It is the logical and
formal interpretation of a word. The second is mainly regarded from four types: 

...... , .........., .........., .........., and .......... . It is the association of lexical items with social factors. Language in such a situation is seen as a metaphorical tool and a .......... phenomenon. It is a device by which individuals can maintain and establish a firm relationship with each other. The third and last type is called .......... meaning. It is the organisation of some detailed information about a subject-matter.
First Semester Exam

Essay question: As opposed to Micro-linguistics, Sociolinguistics has been established as a multidisciplinary field of investigation with a distinct approach. Discuss.

To answer the question, we should deal with each field. It's clear that Micro-linguistics and Sociolinguistics have different dependencies. While the former is isolated and independent, the latter is a multidisciplinary domain. Moreover, both fields do not share the same approaches and methods. Despite the fact that both rely on the same language, they focus on different aspects of language. The interaction of both fields has led to a great deal of research and writing.

To go deeper, we could say that Micro-linguistics has been seen as a free research field, dealing not only with the language itself but also with other factors that may affect it. Language is learning, development, and its possession. Whereas Sociolinguistics is a recent domain and is linked to many other fields.

Finally, Micro-linguistics is based on theoretical approaches that emphasize linguistic generative grammar. While Sociolinguistics is based on the idea that language is not static and it can be modified due to the environment, which people speak it.
Appendix 8

Selected Written Essay ‘B’

University of Djillali Liabes
Faculty of Letters and Arts
Department of English
Full name: Adele Bashir

Module: Linguistics
Academic Year: 2013-2014
Level: Third Year
Group: B
Mark: 62/20

First Semester Exam

Essay question: As opposed to Micro-linguistics, Sociolinguistics has been established as a multidisciplinary field of investigation with a distinct approach. Discuss.

The field of linguistics is so broad and encompasses complexity due to the emergence of subfields that are trying to deal with each seemingly separately. Micro-linguistics, Psycholinguistics, and Laban Linguistics, Sociolinguistics, and Neurolinguistics are examples of different fields set up by researchers.

Sociolinguistics in particular is recognized as an independent and free field of study. It investigates how language is used in society and among people. For example, language changes over time, and these changes are influenced by the society. Language use is impacted by the atmosphere, which may shape the way people use it. In all fields, sociolinguistics is independent and finds its place in different sociolinguistic contexts. Each one has its own approach.
Appendix 9

The Analytic Assessing Process

Macrostructural Process

Assessing the relevance of the content
06/20 to 01/20

Microstructural Process

Paragraphing
03/03 to 01/03

Organisation Coherence
02/02 to 0,5/02

Diction
02/02 to 00/02

Grammar
02/02 to 00/02

Mechanics
02/02 to 00/02

Style
03/03 to 0,5/03
The analytic scoring scheme/ 20/20

1. **Relevance: 06/06**
   - a. knowledgeable : 06/06
   - b. Minimally answer the question. 04/06
   - c. Some knowledge of the subject.03/06
   - d. Limited knowledge of the subject 02/06
   - e. Off-topic .01/20

2. **Paragraphing: 03 /03**
   - a. The number of paragraphs signal the ideas expressed. The essay has a discernible beginning middle and end. 03/03
   - b. Somewhat loosely paragraphed. (The number of paragraphs do not signal the ideas expressed.) 02/03
   - c. The essay lacks an overall plan with beginning, middle and end. 01/03

3. **Style: 03 /03**
   - a. Clear and accurate. 03/03
   - b. Comprehensible, but some occasional errors grammar mechanics and diction affect the meaning of the sentence 02/03
   - c. Comprehensible as in (b) plus some ambiguous words or sentences.01/03
   - d. Chaotic due to frequent errors and ambiguity of some words and sentences. 0, 5/03

4. **Organization /Coherence: 02 /02**
   - a. Ideas cogently organised and logically connected with clear transitions. 02/02
   - b. Somewhat loosely organised but main ideas stand out. 01/02
   - c. Ideas disconnected. Some difficulties to understand the meaning of the answer. 0.5/02

5. **Diction: 02/02**
   - a. Correct choice of words. 02/02
   - b. (-0, 25 / 02) for each wrong word.

6. **Grammar: 02 / 02**
   - a. Accurate sentences (02 / 02)
   - b. (-0, 25) for each grammatical error.

7. **Mechanics: 02 /02**
   - a. Effective use the graphic Convention: spelling, punctuation, capitalisation. 02 / 02
   - b. (-0, 5) for each error
Appendix 11

The Analytical Assessing Grid

<table>
<thead>
<tr>
<th>Students</th>
<th>Written paper ‘A’</th>
<th>Written paper ‘B’</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relevance</td>
<td>06/06</td>
<td></td>
</tr>
<tr>
<td>Paragraphing</td>
<td>03/03</td>
<td></td>
</tr>
<tr>
<td>Coherence/</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organisation</td>
<td>02/02</td>
<td></td>
</tr>
<tr>
<td>Diction</td>
<td>03/03</td>
<td></td>
</tr>
<tr>
<td>Mechanics</td>
<td>02/02</td>
<td></td>
</tr>
<tr>
<td>Grammar</td>
<td>02/02</td>
<td></td>
</tr>
<tr>
<td>Style</td>
<td>02/02</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>20/20</td>
<td></td>
</tr>
</tbody>
</table>
Appendix 12

The Essays Samples Analytical Assessment

<table>
<thead>
<tr>
<th>Students</th>
<th>Written paper ‘A’</th>
<th>Written paper ‘B’</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relevance</td>
<td>06/06 03</td>
<td>02</td>
</tr>
<tr>
<td>Paragraphing</td>
<td>03/03 02</td>
<td>01</td>
</tr>
<tr>
<td>Coherence/ Organisation</td>
<td>02/02 1.50</td>
<td>01</td>
</tr>
<tr>
<td>Diction</td>
<td>02/02 1.50</td>
<td>0.5</td>
</tr>
<tr>
<td>Mechanics</td>
<td>02/02 1.50</td>
<td>01</td>
</tr>
<tr>
<td>Grammar</td>
<td>02/02 1.50</td>
<td>01</td>
</tr>
<tr>
<td>Style</td>
<td>02/02 1.50</td>
<td>01</td>
</tr>
<tr>
<td>Total</td>
<td>20/20 12.50</td>
<td>07.50</td>
</tr>
</tbody>
</table>
Appendix 13

The Teachers’ Interview

Such interview is meant to decipher the teachers’ attitudes and preferences towards the written essays’ assessment procedure.

1- Do you prefer to use the essay as an assessing tool?

2-Which methods do you usually apply in assessing your student’s essays?
   a- Error-count method
   b- Multiple assessing method
   c- Analytic assessing method
   d- Holistic method

3-What does an essay at each of the following scales look like?
   a- Excellent
   b- Very good
   c- Good
   d- High average
   e- Low average
   f- Weak
   g- Very weak
   h- Off-topic

4-Which marks do you prefer to provide to the following essays’ scale?

   Excellent
   Very good
   Good
   High average
   Low average
   Weak
   Very weak
   Off-topic

5-Which of the following essay components do you prefer to rely on in your assessing procedure?
   a- Neatness of the paper.
   b- Content: interest/ originality of ideas.
c- Form: effectiveness of expression

d-Grammar/Syntax

e-Organisation of the ideas

f-Mechanics

g-Handwriting

h-Style

i-Fluency

g-Diction (word choice)

k-Rhetorical matters (organisation, cohesion, unity...)

l-Relevance of the content

6- Do you provide your students’ with any kind of written feedback?
Glossary
**Glossary**

**Assessment:** The process of evaluating and measure an individual’s achievement; typically done with assessment tools such as assignments or examinations.

**Assignment:** A piece of writing that a writer has been assigned to write by an editor or publisher for a pre-determined fee.

**Brainstorming:** Typically used in writing, but is any activity where individuals general ideas related to a topic or task; done in either groups or individually with no restriction on quality of ideas. Once ideas are generation, they are they evaluated and a decision about which to pursue is made.

**Case study:** An investigation format focusing on a specific group, setting, and time period with the aim of studying and clarifying a unique feature(s) of the situation.

**Competencies:** An individual’s abilities as they relate to knowledge, understanding, and skills; see also minimal competence.

**Context:** The setting where teaching and learning occur; identifying context involves noting social, geographical, political, and other factors related to the learning environment.

**Continuous assessment:** Ongoing evaluation of work during a course in which the scores earned count toward the final evaluation.

**Curriculum:** Broadly understood as the subjects and materials to be taught by an educational institution; typically it is listed as a set of subjects, but also may include the learning experiences, skills, and abilities students are expected to learn.

**Diagnostic test:** A test to diagnose or discover what language students know and what they need to develop to improve their language abilities; may be used before a course of study and combined with placement test. It is also an examination used to determine students’ current level of knowledge or skill to identify what course level they should be placed in or whether remediation is required.
**Diction:** In its original, primary meaning refers to the writer's or the speaker's distinctive vocabulary choices and style of expression. A secondary, common meaning of "diction" is more precisely the art of speaking clearly so that each word is clearly heard and understood to its fullest complexity and extremity. This secondary sense concerns pronunciation and tone, rather than word choice and style.

**Draft:** Any preliminary version of a written work. An author may write dozens of drafts which are revised to form the final work, or he or she may write only one, with few or no revisions.

**Edit:** To review a piece of writing to correct grammatical, spelling, or factual errors. Editing often includes shortening or lengthening of a piece of writing to fit an available space before publication.

**Evaluation:** Process of assessing work completed by an individual, group, or institution with the aim of determining whether the individual, group, or institution has meet predetermined standards.

**Facilitator:** Individual who assists others in a learning process but does not act as the primary source of knowledge; the facilitator acts as a guide in during individual or group learning activities.

**Feedback:** Responses provided to an individual while completing a task that is intended to guide the individual to a desired end.

**Formative assessment:** Assessment used to identify an individual’s current strengths and weaknesses relative to a knowledge or skill with the intention of improving one’s knowledge or skill. Observations which allow one to determine the degree to which students know or are able to do a given learning task, and which identifies the part of the task that the student does not know or is unable to do. Outcomes suggest future steps for teaching and learning.
Holistic Assessment: Assigning a single score based on an overall assessment of performance rather than by scoring or analyzing dimensions individually. The product is considered to be more than the sum of its parts and so the quality of a final product or performance is evaluated rather than the process or dimension of performance. A holistic scoring rubric might combine a number of elements on a single scale. Focused holistic scoring may be used to evaluate a limited portion of a learner's performance.

Objective tests: Test based on answers that require students demonstrate a knowledge or skill exactly with no opportunity for judgment by the evaluator.

Peer assessment: Assessment completed by other students or member in the same discipline.

Peer correction: Also known as peer review, peer editing, or peer feedback; in writing, an activity whereby students help each other with the editing of a composition by giving each other feedback, making comments or suggestions; can be done in pairs or small groups.

Performance criteria: Written standards used by an evaluator to judge whether an individual can perform a skill or has demonstrated knowledge.

Reliability: The characteristic that same or similar results can be obtained through repeated experiments or tests.

Remedial teaching: Education activities aimed at removing deficiencies in knowledge or skills.

Revising: Making changes to improve the writing.

Run-on Sentence: Two or more sentences in a paragraph without appropriate punctuation or connecting words.

Self assessment: Assessment completed by the learner him/herself to evaluate his/her own performance, strengths and weaknesses.

Sentence Fragment: It is a sentence that is missing the subject, the verb, or both.
Subjective Test: A test in which the impression or opinion of the assessor determines the score or evaluation of performance. It is a test which answer cannot be known or prescribed in advance.

Summative assessment: Assessment typically completed at the end of a learning period with the aim of providing a final evaluation of individual's mastery of knowledge or a skill. It is a kind of evaluation usually done at the conclusion of a unit of instruction, an activity or plan to determine and judge students’ skills and knowledge, or the effectiveness of a plan or an activity. Outcomes are the culmination of a teaching/learning process for a unit, a subject, or a year’s study

Style: The manner of expression of a particular writer, produced by choice of words, grammatical structures, use of literary devices, and all the possible parts of language use

Validity: The degree to which an investigation accurately assesses the specific idea a researcher is investigating.
Résumé / La conception de la méthode la plus efficace pour évaluer les capacités de l’écrit des apprenants d’Anglais comme langue étrangère est devenue une question controversée. Plusieurs recherches ont tenté une approche d’évaluation efficace et fiable pour interférer de manière à stimuler les compétences de l’écrit chez les apprenants. Or, l’enseignant et l’apprenant sont toujours confus sur le sujet. Le premier devient enchevêtré à regarder des caractéristiques identiques inappropriées qui se répètent chaque fois qu’il évalue un essaie. Alors, le dernier remarque souvent que sa progression est lente avec des erreurs similaires qui se reproduisent. Par conséquent, la présente contribution est une humble tentative d’améliorer la fiabilité et l’objectivité de l’évaluation de l’écrit en procréant une formule d’évaluation analytique. Une telle procédure peut examiner tous les aspects des aptitudes de l’écrit chez les apprenants. Par ailleurs, l’étude attire l’attention sur la possibilité d’associer la technologie informatique dans l’amélioration de la fiabilité de l’évaluation analytique. L’enquête menée a démontré que la plupart des enseignants évaluent l’écrit de manière holistique, pour cette raison, aucune rétroaction efficace ne peut être conçue et certainement aucune amélioration ne peut avoir lieu. D’autre part, l’intégration d’une procédure d’évaluation analytique convertie en programme d’informatique peut augmenter la fiabilité de l’évaluation, améliorer l’efficacité de la rétroaction et sans doute augmenter les performances d’écrit des apprenants.

Mots clés: L’évaluation, la fiabilité, la validité, l’objectivité, l’écrit, l’évaluation holistiques, l’évaluation analytique, l’évaluation automatique, compétences de l’écrit

Summary/ Designing the most successful method to assess the EFL learners’ writing abilities has become a controversial issue. An enormous amount of researches have attempted an effective and reliable writing assessment approach to interfere so as to boost the learners’ writing proficiency. Yet, both the EFL teacher and the EFL learner are still confused about the subject matter. The former becomes troubled at watching identical inappropriate features happening yet again. While, the latter frequently notices that his progress is slow with similar errors recurring. Consequently, the present contribution is a humble attempt to improve the reliability and objectivity of the teacher’s assessment to the written language through generating an analytic assessing formula. Such a procedure may scrutinize every aspect of the learner’s writing aptitudes. Thus, the teacher may have an opportunity to identify his learners’ hindrances and to design the appropriate feedback. Besides, the study draws attention to the option of involving the computer technology in enhancing the analytic assessment reliability. The investigation has shown that most EFL teachers are assessing their learners’ writing holistically, for that reason no effective feedback could be designed and certainly no improvement could take place. On the other hand, the integration of an analytic assessing procedure converted into computer programme may increase assessment reliability, enhance the feedback effectiveness and without doubt boost the learners’ writing performance.

Key words: Assessment, reliability, validity, objectivity, writing, holistic assessment, analytic assessment, automatic assessment, writing skills