



Ministry of Higher Education and Scientific Research

Constantine 1 University

Faculty of Letters and Languages

Department of English



**Gender-based Evaluation of Learners' Academic Achievements:
The Case of Brevet and Baccalaureate Examinees in the Wilaya
of Bordj Bou Arreridj, Algeria**

**Thesis submitted to the Department of English in candidacy for the degree
of “Doctorate es-science” in Applied Linguistics**

Submitted By:

Mr Riadh BOUKHETALA

Supervised by:

Prof. Riad BELOUAHEM

Board of Examiners

Chairperson	Prof. Nacif LABED	Professor	Constantine 1 University
Supervisor	Prof. Riad BELOUAHEM	Professor	Constantine 1 University
Member	Prof. Youcef BEGHOUL	Professor	Constantine 1 University
Member	Prof. Said KESKES	Professor	Setif 2 University
Member	Prof. Hacene HAMADA	Professor	Teachers' Training School, Constantine
Member	Prof. Nadir KAOULI	Professor	Batna 2 University

Dedication

To my dear parents, my precious wife -*Asma*-, and my two young men

– *Mouhamed Firas and Ali Assil*–

To all those, who supported me and believed in me throughout my
life: family, friends, and teachers...

A special dedication to this work is due to the souls of my friends
Mohamed Tahar Asses and *Rachid Hannachi* who passed away while
they were left with a few steps before the accomplishment of their
PhD theses ... May God bless their souls!

Acknowledgements

First and foremost, I would like to thank my supervisor Prof. Riad Belouahem for his support, encouragement and patience along the way leading to the achievement of this research work. He has been a teacher, a supervisor, and a mentor. My thanks are also due to the board of examiners each one by their title and name for the time, effort, engagement, and contribution in polishing this thesis to give it more value and raise its potential contribution to the field of research.

I am also thankful for the hospitality, help and thoughtful insights provided by Prof. David Putwain and Dr Jo Frankham at the Liverpool John Moores University (LJMU), the provision of data by Dr Ahmed Belemrabta (Head of School Guidance and Orientation Centre of the Wilaya of Bordj Bou Arreridj), and last but not least to the support and advice given by Dr Mouloud Ait Aissa, Dr Mohamed Saadaoui, Dr Najat Khenioui, and Dr Sid Ali Selama.

ABSTRACT

This study aims to find out whether there is a significant difference between male and female learners in Algerian middle and secondary schools' performances in the BEM and BAC English language examinations. It also aims at studying the degree of homogeneity of the BEM and BAC exam papers with the textbooks' contents in terms of topics and activity types. All this would lead to conclusions about whether gender is a major indicator of performance or whether there are other possible factors that influence learners' performances regardless of their gender. A descriptive, cross-sectional research design was adopted using two main research tools and procedures. First, a quantitative and qualitative content-based analysis of the 'On the Move' and 'New Prospects' textbooks was conducted based on the topics treated in each unit/file in addition to the activity/task types. The same analysis was applied to the BEM and BAC exams (scientific and foreign languages streams only) for the years 2014 and 2019. Then, a statistical analysis of male and female BEM and BAC examinees' results in the Wilaya of Bordj Bou Arreridj in the years 2014 and 2019 was carried out. Descriptive statistics of the two populations of examinees were discussed on a comparative basis; whereas, inferential statistics using two sample z-tests considered only samples of male and female examinees (from both the 2014 and 2019 populations) that were randomly selected by using the SPSS software. The results have shown that there is a high degree of homogeneity between the BEM and BAC English language exam papers and the instructional textbooks used by Algerian learners in terms of topics and task/activity types. The statistical analysis revealed that there is no difference between male and female learners' means of scores in the foreign languages BAC examinees' results in both years. Moreover, although this was not correct by analysing the two sample z-tests results for the BEM and the scientific stream BAC examinees' scores, the effect size was very small (BEM: $d=0.07/0.1$, BAC scientific stream: $d=0.07/0.04$ in 2014 and 2019 respectively), which means that the magnitude of the difference is trivial and can be neglected. These findings lead to assume that *the gender variable taken alone does not determine the possibility and the rate of students' success on tests and exams*. Hence, the researcher referred to other possible extraneous variables that, according to the analysis findings, seem to be relevant factors influencing academic achievements.

List of Tables

Table	Page
Table 1.1. Stages, decision-making roles and products in curriculum development.....	27
Table 2.1. Contextual Factors Involved in Planning a Language Programme.....	41
Table 2.2. Textbook Analysis at Three Levels Based on Littlejohn (1998).....	54
Table 2.3. The Categories and Types of Assessment	58
Table 2.4. Differences between Norm-Referenced and Criterion-Referenced Tests.....	61
Table 2.5. Matching Tests to Decision Purposes.....	68
Table 3.1. Simple Frequency Distribution of 20 Students' Class Scores	96
Table 3.2. Grouped Frequency Distribution of 20 Students' Class Scores.....	97
Table 4.1. Activity / Task Types in BEM 2014 & BEM 2019 English Language Examination.....	144
Table 4.2. Duration and Coefficient of the English Subject in the BAC Exam.....	146
Table 4.3. BAC 2014 English Language Exam Paper Organisation.....	147
Table 4.4 BAC 2019 English Language Exam Paper Organisation.....	147
Table 4.5. Free Writing Topics for BAC 2014 & 2019 (Foreign Languages Stream).....	149
Table 4.6. Activity / Task Types in BAC 2014 & BAC 2019 English Language Examination (Scientific Stream).....	150
Table 4.7. Activity / Task Types in BAC 2014 & BAC 2019 English Language Examination (Foreign Languages Stream).....	152
Table 5.1. Frequency of BEM 2014 Population.....	159
Table 5.2. Frequency of BEM 2019 Population.....	159
Table 5.3. BEM 2014 Population Descriptive Statistics	160
Table 5.4. BEM 2019 Population Descriptive Statistics	162

Table 5.5. z-test: Two Sample for Means BEM 2014.....	164
Table 5.6. z-test: Two Sample for Means BEM 2019.....	166
Table 5.7. Frequency of BAC 2014 Population.....	168
Table 5.8. Frequency of BAC 2019 Population.....	168
Table 5.9. BAC 2014 (Scientific Stream) Population Descriptive Statistics	169
Table 5.10. BAC 2019 (Scientific Stream) Population Descriptive Statistics	170
Table 5.11. BAC 2014 (Foreign Languages Stream) Population Descriptive Statistics.....	172
Table 5.12. BAC 2019 (Foreign Languages Stream) Population Descriptive Statistics.....	173
Table 5.13. z-test: Two Sample for Means BAC 2014 (Scientific Stream).....	176
Table 5.14. z-test: Two Sample for Means BAC 2019 (Scientific Stream).....	178
Table 5.15. z-test: Two Sample for Means BAC 2014 (Foreign Languages Stream).....	179
Table 5.16. z-test: Two Sample for Means BAC 2019 (Foreign Languages Stream).....	180

List of Figures

Figure	Page
Figure 2.1. The Framework of Language Teaching	40
Figure 2.2. Levels of Analysis of Language Teaching Materials	53
Figure 2.3. Relationships among Measurement, Tests and Evaluation	57
Figure 2.4. Language-Test Task Characteristics.....	66
Figure 3.1. Summary of the Structural Changes from the Original Framework to the Revision.....	83
Figure 3.2. Model of Behaviour.....	84
Figure 3.3. The New Taxonomy.....	85
Figure 3.4 Histogram Based on a Grouped Frequency Distribution	99
Figure 3.5. Grouped Frequency Distribution with a Frequency Polygon.....	100
Figure 3.6. A Smooth Curve.....	101
Figure 3.7. Importance of the Median to Represent Central Tendency in Skewed Distributions.....	104
Figure 3.8. Quartered Distribution.....	106
Figure 3.9. Per cent of Cases under Portions of the Normal Curve.....	110
Figure 5.1. English Mark * Gender Cross tabulation Chart (BEM 2014).....	161
Figure 5.2. English Mark * Gender Cross tabulation Chart (BEM 2019).....	163
Figure 5.3. English Mark * Gender Cross tabulation Chart (BAC 2014, Scientific stream).....	171
Figure 5.4. English Mark * Gender Cross tabulation Chart (BAC 2019, Scientific stream).....	171
Figure 5.5. English Mark * Gender Cross tabulation Chart (BAC 2014, Foreign Languages stream).....	174
Figure 5.6. English Mark * Gender Cross tabulation Chart (BAC 2019, Foreign Languages stream).....	175

List of Acronyms and Abbreviations

BEM: Brevet Exam

BAC: Baccalaureate Exam

CLT: Communicative Language Teaching

ELT: English Language Teaching

***f*:** Frequency

H_0 : Null Hypothesis

H_1 : Research / Alternative Hypothesis

***i*:** Interval

IELTS: International English Language Testing System

IQR: Interquartile Range

MBOE 1: My Book of English Year 1

MBOE 2: My Book of English Year 2

MBOE 3: My Book of English Year 3

MBOE 4: My Book of English Year 4

MCQs: Multiple-Choice Questions

MDN: Median

N: Population/Sample Number

N_1 : Male Population/Sample Number

N_2 : Female Population/Sample Number

NP: New Prospects

OTM: On the Move

Q: Quartile

R: Range

s^2 : Sample Variance

SD / σ : Standard Deviation

SE1: Secondary Education year 1

SE 2: Secondary Education year 2

SE3: Secondary Education year 3

SIQR: Semi-interquartile Range

Skew₁: Skewness for Male Population Distribution

Skew₂: Skewness for Female Population Distribution

SPSS: Statistical Package for Social Scientists

TBLT: Task-Based Learning and Teaching

TLU: Target Language Use

TOEFL: Test of English as a Foreign Language

μ : Population Mean

μ_1 : Male Population Mean

μ_2 : Female Population Mean

$\bar{x}$: Sample Mean

X : Scores

σ^2 : Population Variance

Σ : Sum

Table of Contents

Dedication	I
Acknowledgements	II
ABSTRACT	III
List of Tables.....	IV
List of Figures	VI
List of Acronyms and Abbreviations	VII
General Introduction	1
1. Statement of the Problem.....	1
2. Aim of the Study.....	3
3. Research Questions.....	2
4. Research Hypotheses.....	3
4. Research Methodology	3
5. Structure of the Thesis	5
Chapter One: Males, Females and the Algerian Curriculum	
Introduction	9
Section One: Gender Differences and Similarities	9
1.1.1. Definition of Gender.....	9
1.1.2. Gender Differences.....	10
1.1.2.1. Biological Differences.....	10
1.1.2.2. Physical Differences	11
1.1.2.3. Cognitive Differences.....	12
1.1.2.4. Psychological Differences	20
1.1.2.5. Sociological Differences	20
1.1.3. Gender Similarities.....	24
Section Two: The Place of Gender within the Algerian Curriculum	25
1.2.1. Definition of Curriculum.....	25
1.2.2. Language Curriculum Development	25
1.2.3. The Algerian Curriculum: A Reflection on the Educational Policy.....	27
1.2.4. The English Language Curriculum in Algeria	30
1.2.5. Learners' Gender within the Algerian Curriculum.....	34
Conclusion.....	35

Chapter Two: Syllabus Design, Teaching/Learning Materials and Assessment & Evaluation: Basic Components of the Curriculum

Introduction	37
2.1. Syllabus Design	37
2.1.1. Definition of a ‘Syllabus’	37
2.1.2. Syllabus Design vs. Methodology	38
2.1.3. Syllabus Design and Teaching Materials	39
2.1.3.1. Defining ‘Materials’	39
2.1.3.2. The Framework of Materials and Methods in ELT	40
2.2. Assessment/Evaluation	57
2.2.1. Types of Assessment	58
2.2.1.1. Informal/Formal Assessment.....	59
2.2.1.2. Formative/Summative Assessment	59
2.2.1.3. Norm-referenced/Criterion-referenced Assessment	60
2.2.1.4. Classroom-based/Large-scale Assessment	62
2.2.2. Test Types	62
2.2.2.1. Placement Test	62
2.2.2.2. Proficiency Test.....	63
2.2.2.3. Diagnostic Test.....	63
2.2.2.4. Achievement Test.....	64
2.2.3. Framework of Language Test Task’s Characteristics	65
2.2.4. Matching Tests to Decision Purposes.....	67
2.2.5. Cardinal Criteria for ‘Testing a Test’	69
2.2.5.1. Validity	69
2.2.5.2. Reliability	72
2.2.5.3. Practicality	73
2.2.5.4. Equivalency	74
2.2.5.5. Washback	74
Conclusion.....	75

Chapter Three A Focus on Achievement Tests

Introduction	76
3.1. Nature and Features of Achievement Tests	76
3.2. Functions of Achievement Tests.....	77
3.3. Limitations of Achievement Tests	77
3.4. Taxonomies of Educational Objectives and Achievements.....	78
3.4.1. Bloom’s Taxonomy	78
3.4.2. Anderson and Krathwohl's Taxonomy	81
3.4.3. Marzano and Kendall’s New Taxonomy	84
3.5. Achievement Test Planning and Design	85
3.5.1. Setting the Purpose for Testing	85
3.5.2. Developing a Test Plan.....	86
3.5.3. Selection of Item Formats	86
3.5.3.1. Essay Tests	86
3.4.3.2. Objective Tests	88
3.5.4. Quantity of Test Items	93
3.5.5. Distribution of Test Items’ Difficulty Level.....	94
3.6. Describing and Summarising Achievement Test Results	95
3.6.1. Summarizing Data.....	95
3.6.1.1. Frequency Distribution	95
3.6.1.2. Graphing Data	99
3.6.2. Describing and Summarising Score Distributions.....	102
3.6.2.1. Measures of Central Tendency	102
3.6.2.2. Measures of Variability	105
3.6.2.3. Normal Distribution.....	108
3.6.2.4. Standard Scores	110
Conclusion.....	111
Chapter Four: Thematic and Task-Based Analysis of English Language Textbooks and BEM and BAC Exam Papers	
Introduction	112
Section One: Thematic and Task-Based Analysis of ‘On the Move’ and ‘New Prospects’ ..	112
4.1.1. Presentation of the Textbooks	112

4.1.1.1.	On the Move (OTM)	112
4.1.1.2.	New Prospects	115
4.1.2.	Content-based Analysis	117
4.1.2.1.	Topics / Themes	118
4.1.2.2.	Activity/Task-Based Evaluation of “On the Move” and “New Prospects”	138
Section Two: Thematic and Items’ Analysis of English Language BEM and BAC Exam Papers		142
4.2.1.	The BEM English Language Exam	142
4.2.1.1.	Description of the Exam Paper	142
4.2.1.2.	Analysis of the Exam Papers	143
4.2.2.	The Baccalaureate English Language Exam	146
4.2.2.1.	Description of the Exam Papers	146
4.2.2.2.	Analysis of the Exam Paper	148
4.2.3.	Summary of the Findings	154
Conclusion		155
Chapter Five: Gender-Based Analysis and Discussion of Learners’ Academic Achievements in English Language BEM and BAC Exams		
Introduction		156
5.1. Research Methodology		156
5.1.1.	Population and Sampling	156
5.1.2.	Data Analysis Tools	157
5.1.2.1.	z-Tests, t-Tests and the Central Limit Theorem	157
5.1.2.2.	Cohen’s <i>d</i> Effect Size	159
5.2. BEM English Language Examination		159
5.2.1.	Population	159
5.2.2.	Descriptive Statistics and Analysis	161
5.2.2.1.	BEM 2014	161
5.2.2.2.	BEM 2019	163
5.2.3.	z-Tests/Effect Size Results and Interpretation	164
5.2.3.1.	BEM 2014	165
5.2.3.2.	BEM 2019	166

5.3. Baccalaureate English Language Examination.....	168
5.3.1. Population.....	168
5.3.2. Descriptive Statistics and Analysis	170
5.3.2.1. Scientific Stream	170
5.3.2.2. Foreign Languages Stream	173
5.3.3. z Tests/Effect Size Results and Interpretation	177
5.3.3.1. Scientific Stream	177
5.3.3.2. Foreign Languages Stream	180
5.3. Summary of the Findings.....	183
5.3.1. Descriptive Statistics	183
5.3.2. z-Tests.....	184
Conclusion.....	185

Chapter Six: General Discussion, Implications, and Recommendations

Introduction	185
6.1. Exam papers vs. Textbooks: Content-based Comparison.....	185
6.1.1. BEM Exam Papers and ‘On the Move’	185
6.1.2. BAC Exam Papers and ‘New Prospects’	188
6.2. Discussion of the Statistical Data	192
6.2.1. The BEM Exam Statistics	192
6.2.2. The BAC Exam Statistics.....	193
6.3. Summary of the Discussions.....	195
6.4. Recommendations and Implications	198
6.4.1. To the Policy Makers and Curriculum Developers	198
6.4.2. To the Syllabi and Textbook Designers	200
6.4.3. To the Exams’ Designers.....	201
6.5. Limitations of the Study.....	202
6.6. Suggestions for Further Research	202
Conclusion.....	203
General Conclusion.....	204

REFERENCES.....	207
Appendix A: A Schedule for Analysing Tasks (Littlejohn, 2011, p.191).....	215
Appendix B: Sample Tasks' Analysis Sheet (Adapted from Littlejohn (2011, p191).....	216
Appendix C: 'On the Move' Coursebook Adaptation Chart.....	217
Appendix D: Structure and Typology of the BEM Exam Activities	218
Appendix E: Typology of the BAC Exam Activities.....	221
Appendix F: English Language Examination BEM 2014.....	225
Appendix G: English Language Examination BEM 2019	227
Appendix H: English Language Examinations BAC 2014.....	229
Appendix I: English Language Examinations BAC 2019	238

General Introduction

1. Statement of the Problem

The gender gap in educational achievements all over the world has long been a topic of debate within the educational field and it has also attracted the attention of researchers in other fields, the media, and the public as well. Titles of newspaper articles such as '*Academic excellence of female to male turn into a phenomenon*' from El Khabar and '*Females' excellence at studies*' found in El Massa newspaper may orient the general opinion inside and outside the country towards a conviction that male learners are poor achievers and the gap is so wide to become a *phenomenon*. Moreover, since academic achievements are strongly linked to school dropout rates, they are given much importance by the political and educational authorities. The ministry of national education in Algeria has made many changes and reforms within the Algerian educational system starting from 2002. Examples of these reforms are the shift from the fundamental school to the primary and middle school system, in addition to raising the number of years of obligatory schooling from 6 to 9 years to fight learners dropout at an early age. Aiming to achieve this and to enhance the quality of education, it was necessary to reform and renew curricula, syllabuses and textbooks.

In this research, the argument which is put forward is that the topic of the gender gap in Algerian learners' academic achievements should not be studied in isolation but rather, the top-down approach must be adopted. In other words, learners' performances at the end of cycles' exams are the outcome of an educational policy, which is put in some policy documents (curricula). These are, then, manifested in the syllabi of the different school subjects, the instructional textbooks used for the teaching/learning of those subjects, and finally in the tests and exams that result in scores used in the interpretation of the learners' achievements. Hence, the study of the whole chain, in addition to a scientific (statistical) analysis of male and female

learners' scores, is more than necessary to arrive at objective, reasonable, and holistic conclusions about the so-called *gender gap phenomenon*.

2. Aim of the Study

This research aims at reaching conclusions about the real situation concerning the academic achievements of Algerian male and female learners in official English language examinations such as the BEM and BAC exams. The aim behind opting for these exams is the fact that examinees' papers are anonymous and peer-corrected which offers more credibility for the results and achieves a higher degree of objectivity in the research. Since the exams' scores are the outcome of examinees' performance on a test (exam), the analysis of the English language BEM and BAC exam papers and the instructional textbooks was necessary. This is to check their degree of homogeneity in order to ensure that the learners are tested based on the contents they have been dealing with in class and that no extraneous factors -in terms of exam content- may have affected their achievements. Hence, the question of whether *gender*, as a variable, can be regarded as an indicator of learners' achievements in English language tests and exams or not will be answered.

3. Research Questions

Three main research questions are asked in this study:

- 1- Are the *testing* and the *instructional* English language materials (i.e. exams and textbooks) designed for Algerian middle and secondary school learners, sitting for the BEM and BAC exams, homogeneous in terms of topics and task/activity types?
- 2- Is there a gender gap in Algerian middle and secondary school learners' academic achievements in BEM and BAC exams? If yes, to which extent?
- 3- Based on the above questions, can gender be considered a primary indicator of academic achievements?

4. Research Hypotheses

The null hypotheses set for the previously asked questions are the following:

- a. There is a high degree of homogeneity between the BEM and BAC English language exam papers and the instructional textbooks used by Algerian learners in terms of topics and task/activity types.
- b. There is no difference between male and female learners' means in the BEM and BAC English language exams.
- c. Gender, taken alone, does not determine the possibility and the rate of students' success on tests and exams.

The alternative hypotheses are set as follows:

- a. BEM and BAC English language exam papers' contents do not reflect the contents of the instructional textbooks used by Algerian learners in terms of topics and task/activity types.
- b. There is a significant difference between male and female learners' means in the BEM and BAC English language exams.
- c. Gender is a variable that determines the possibility and the rate of students' success on tests and exams.

5. Research Methodology

a. Research Design

The present research is both descriptive and cross-sectional in nature. The cross-sectional design involves both a descriptive and an analytical study. To overcome the drawback of cross-sectional studies, which study only a single point in time, the present research targeted four different populations at two different points in time with a period of five years (2014-2019) between the first and the second exams that these populations sat for. A mixed-method approach

is used in this research. Both quantitative and qualitative methods are employed for analysing data.

b. Research Tools and Procedures

- **Content analysis:**

The fourth-year middle school English textbook ‘On the Move’ and the third-year secondary school English textbook ‘New Prospects’ were analysed based on the topics/themes they discuss and the types of activities/tasks they suggest as learning support. The same thing was done with the BEM and BAC exam papers. This analysis was conducted on a qualitative basis (topics) and a quantitative one (the classification of activities and tasks using Littlejohn’s Schedule for Analysing Tasks (2011) that was adapted by including Bloom’s taxonomy’s levels of cognitive skills). The aim here is to answer the first research question and therefore accept or refute the first null hypothesis.

- **Descriptive and inferential statistics:**

Kubiszyn and Borich (2013) consider that “Descriptive statistics are simply numbers, for example, percentages, numerals, fractions, and decimals. These numbers are used to describe or summarize a larger body of numbers” (p.263). In this research, a set of descriptive statistics is adopted including cross-tabulation charts of male and female scores, measures of central tendency (mode, median, and mean), measures of variability (standard deviation, variance, and range) in addition to an analysis of the male and female scores’ distribution in terms of skewness and kurtosis.

Inferential statistics, on the other hand, tries to reach conclusions that extend beyond the immediate data alone. In this regard, two sample z-tests were used to compare the mean scores of male and female learners to find out whether there is a significant difference between the two groups’ means or not. The effect size is also calculated to measure the magnitude of the

difference between the two groups. In our research, we opted for Cohen's d as a measure for this. All this aims to either confirm or reject the second null hypothesis.

c. Population and Sampling

Male and female learners who sat for the 2014 and 2019 BEM and BAC exams in the Wilaya of Bordj Bou Arreridj represent the populations of the present study. Concerning the BAC exam, the study focused on learners from two major streams, which are the *Scientific* the *Foreign languages*. The choice of these populations was not made randomly but purposefully due to some special conditions and circumstances that characterized the school years and the official exams in the years between 2014 and 2019 (e.g. Exam papers' leak on the social media before the exam date). These, from the researcher's point of view, could bias the study's results and made him opt for the aforementioned exams.

Random samples from these populations were selected with the help of the SPSS software to conduct the two sample z -tests. Approximately, 20% of the whole population of each exam (including male and female learners) was selected as a sample.

6. Structure of the Thesis

This thesis consists of six chapters, among which the first three chapters are theoretical in nature and they adopt the top-down approach, or in other words, they start from general to specific starting with the curriculum and ending with achievement tests. The last three chapters deal with the analysis of the textbooks and the exam papers, and the analysis of the descriptive and inferential statistics about male and female learners' scores in the chosen exams. In the sixth and last chapter, a general discussion of the two previous chapters' results is presented together with some educational and pedagogical implications of the present research.

Chapter one, which deals with gender and the Algerian curriculum consists of two sections. The first one discusses gender as a social construct considered by some researchers to be different from biological sex. However, in this study, as in much other literature, the terms *sex* and *gender* are used interchangeably to refer to males and females. The focus is, then, put on the differences and similarities between the two sexes. Section two, on the other hand, deals with the basic elements of curriculum development and sheds light on the policy documents that are used as a reference for the design of the curricula for the different school subjects in the Algerian schools and the English language curriculum in particular. It ends up with a presentation of the place and status of learners' gender in these policy documents.

Chapter two entitled - Syllabus Design, Teaching/Learning Materials and Assessment & Evaluation: Basic Components of the Curriculum - reviews the literature related to syllabus design, materials/textbooks' analysis and evaluation, and assessment, testing and evaluation. Hence, a distinction is made between the narrow and broad views of syllabus design paving the way to discuss the relationship between syllabus design and materials design through the presentation of *the framework of language teaching*. These materials, and textbooks, in particular, being teaching-learning materials, can be analysed and/or evaluated based on different methods, from different perspectives and for various purposes. The chapter concludes by dealing with the types of assessments, and types of tests in addition to the characteristics of a 'good' test.

Chapter three, as its title – A Focus on Achievement Tests- implies, can be described as a zoomed picture of achievement tests that are one of the test types discussed in the previous chapter. The chapter sheds light on the advantages and limitations of these tests, presents the steps of achievement tests' planning and design, and links them with the educational objectives that need to be reached by referring to some taxonomies of educational objectives that are found in the literature. Some theoretical background about a set of basic notions and items in descriptive and inferential statistics that are essential in describing and summarising achievement tests' results is dealt with at the end of the chapter.

Chapter four presents the results of the analysis of the 'On the Move' and 'New Prospects' English language textbooks for the fourth year of middle school and third year of secondary school respectively. It also analyses the contents of the English language exam papers for the 2014 and 2019 BEM and BAC exams, which are the official end-of-cycle exams (achievement tests) in Algerian middle and secondary school cycles. Both qualitative analysis (themes) and quantitative analysis (task/activity types) methods were applied. A summary of the findings is provided at the end of the chapter.

Chapter five is a purely statistical one. It starts by explaining some notions in inferential statistics that are essential to reach meaningful results and to draw significant conclusions. The chapter then proceeds by presenting and discussing the results of the descriptive statistics concerning male and female learners' scores, and the results of the two sample z-tests to compare male and female learners' score means in the selected English language BEM and BAC exams. A summary of the findings of both types of statistical data (descriptive and inferential) is found at the end of the chapter.

In chapter six, a general discussion of the results of chapters four and five is provided. Moreover, a trial to make links between male and female scores and score means' results, on the one hand, and the exams' characteristics, forms and contents, on the other hand, is made. The links extend to the textbooks being the instructional materials used in the teaching-learning process that precedes those exams, and to the courses' syllabi and the English language curricula for both school levels, in particular. This is put in the form of pedagogical and educational implications, which are followed by some recommendations to the policymakers and educational authorities, syllabus and textbook designers, as well as exam papers' writer

Chapter One

Males, Females and the Algerian Curriculum

Table of Contents

Introduction	9
Section One: Gender Differences and Similarities	9
1.1.1. Definition of Gender.....	9
1.1.2. Gender Differences.....	10
1.1.2.1. Biological Differences.....	10
1.1.2.2. Physical Differences	11
1.1.2.3. Cognitive Differences.....	12
1.1.2.4. Psychological Differences	20
1.1.2.5. Sociological Differences	20
1.1.3. Gender Similarities.....	24
Section Two: The Place of Gender within the Algerian Curriculum	25
1.2.1. Definition of Curriculum.....	25
1.2.2. Language Curriculum Development	25
1.2.3. The Algerian Curriculum: A Reflection on the Educational Policy.....	27
1.2.4. The English Language Curriculum in Algeria	30
1.2.5. Learners' Gender within the Algerian Curriculum.....	34
Conclusion.....	35

Introduction

The curriculum in any educational system is the manifestation of that system's philosophy, ideology, aims as well as educational policy. This latter, reflects the vision of the policy/decision-makers vis-à-vis the educational issues, the targeted ends of education, the role of the school, the status and mission of the various actors in the educational field...etc. Moreover, the learner is certainly the principle and basic component of the educational system and the one by which and for which the whole system is meant to exist and operate. One of the issues that educational policy-makers tend to focus on in the process of curriculum design is taking into consideration the learners' sex/gender as many pieces of research were conducted to study the differences (and to a lower extent the similarities) between girls and boys. In this chapter, we start with a discussion of gender differences and similarities and then a move is made to deal with the curriculum in general; its definition and steps of development, and in particular to discuss the Algerian curriculum and the importance it attaches to gender.

Section One: Gender Differences and Similarities

1.1.1. Definition of Gender

Gender and sex are terms used interchangeably, in many cases, to refer to the same thing. Nevertheless, people working in the field of gender, in general, make a distinction between the two terms.

Eckert and McConnell-Ginnet (2003) point out that: "gender is the social elaboration of biological sex. Gender builds on biological sex, it exaggerates biological difference and, indeed, it carries biological difference into domains in which it is completely irrelevant."(p.10)

APA sixth edition (2010) states that: "... gender refers to role, not biological sex, and is cultural." (p.73)

1.1.2. Gender Differences

1.1.2.1. Biological Differences

Most studies carried out about the human species since their early formation showed that chromosomes and hormones have a big deal in determining the sex of the newborn baby. Almost all of us have in our AND 46 chromosomes: Half of them from the father (23 chromosomes) and the other half from the mother. Two types of chromosomes exist; X and Y. All the 45 first chromosomes are X ones. The last chromosome from the father makes the difference. A male baby would come to life in case it is a Y chromosome (XY baby); otherwise, in the case of an X one, a female baby is to be born (XX baby). However, tests have revealed variations in chromosomes, including XXY, XYY, and XXX. Anatomically, males and females have different reproductive organs: a penis, testicles, and scrotum for males, and a vagina, uterus, and ovaries for females. Other anatomical differences are represented by the development of breasts among females and the presence of a menstrual cycle. Meyer (2010) explains the categorisation of males and females by saying:

Early in development, the gonadal tissue is undifferentiated. This means that if a Y chromosome is present it stimulates the production of the Testis Determination Factor (TDF) protein which causes testes to develop. In the absence of this protein, ovaries develop. In other words, all embryos start as female and will develop as such unless TDF is present to create testes. (pp. 29-30)

Richardson (2012), in her article: *Sexing the X: How the X Became the "Female Chromosome"*, tried to defend the feminist view that the extra X chromosome is a gift to females and a sign of superiority as she bases her argument on some studies such as Berman (1921), Montagu (1953), and Mundell (2007). She also reports from Mundell his idea that "nature hands women extra ammunition right from the start. The reason, according to geneticists: Is females are gifted with two copies of the powerful X chromosome, while males

are born with only one X, plus the relatively weak Y chromosome.” (p.916). She, also, refers to the work done by Migeon (2007), in which she considers that the extra genetic material found in the extra X chromosomes in females might explain some differences between the sexes and the reason why girls are more attentive and outperform boys in school from the right beginning. (pp. 916-917)

Hormones also have a role in deciding the sex of the newborn baby. Normal people traditionally considered as being male or female have bodies which produce hormone levels adequate to their sex i.e. male bodies produce “male hormones” and female bodies produce “female hormones”.

1.1.2.2.Physical Differences

Throughout their life span, human beings witness a lot of body changes and modifications. From the early years of birth, these changes start to shape the bodies of males and females. The following is a demonstration of humans` physical development stages and the characteristics of each stage for both sexes as summarised by Becker (1986):

- 1- The early years:** the only differences, at this period, between male and female bodies are noticed in their genitals i.e. penis and testicles for males and vagina and ovary for females. The hair on female babies` heads seems to grow faster than that of male ones. The variables of interest in physical growth, according to Becker at this particular period, are nutrients, health, genes, and time.
- 2- The middle years (6-12):** At the age of 6, the average person is 4 feet tall, while at the age of 12, the average is 5 feet. According to Becker, girls are usually shorter at the age of 6, equal at 9 and taller than boys at 12. Brain size does not differ greatly when comparing its size at an early age with that of age 20; it represents 90% of its size at age 5 compared with adulthood size and 95% at age 10.

3- Puberty and adolescence: It is the period when basic and distinguishing differences between the sexes reach their peak. Adolescence refers to the period that signals the beginning of puberty. The physical changes that characterise this period occur between the age of 9-16 for girls and 10-17 for boys. Becker stresses two main points about these changes: “(1) they occur a year or two earlier for girls on the average, and (2) within sex groups the range in age of onset is over two years.” (p.317). A point of interest concerning these changes is that girls have the capacity of being pregnant at the age of 14; whereas, boys can produce mature sperm at the age of 15 on average.

Brannon (2017) considers that physical differences represent the largest differences between the sexes. She states that:

This difference relates to size and muscle mass, with men being significantly larger and stronger than women. The differences among individuals are also large; some women are stronger than other women, and some men are stronger than others. However, gender differences are larger than individual differences, making gender a good predictor of strength. (p. 510)

These differences in physical strength, Brannon believes, also explain differences in gender roles, activities and occupations in society.

1.1.2.3.Cognitive Differences

Women and men are believed to differ in terms of cognitive abilities and this is shown in many areas. It is said, for example, that men outperform women in mathematical and spatial abilities, and that women outperform men in verbal abilities. Statt (2003) explains *cognitive abilities* as “A general term which includes all the mental processes by which people become aware of, and gain knowledge about, the world” (p.26); however, Richardson (1997) says that “...the psychological research exploring gender differences has focused on particular domains of cognitive functioning: mathematical, spatial, and verbal abilities” (p. 3).

The difference in the cognitive abilities between males and females was thought to be caused by the size and weight of the brain. Hence, men are intellectually better than women because their brains are bigger and heavier. However, this assumption was refuted and criticised as this would lead us to think that human beings, men, in particular, are intellectually inferior to some animals whose brains are bigger and heavier.

The differences and similarities in cognitive abilities between males and females are presented in detail under the following subtitles.

1.1.2.3.1. Intelligence

“Which sex is smarter than the other?” is a question asked by many researchers such as Jensen (1998), who conducted surveys and tests for the sake of coming up with convincing answers. Richardson (2012), when explaining the relationship between the X chromosome and the degree of intelligence of the sexes, says that “Although the average IQ of men and women is equal, men are more frequently found at both extremes of intelligence” (p.918). Halpern (2012) discusses the question previously asked by saying that the problem with it is that it is built on the prior “...assumption that *there is a smarter sex*” (p.98).

Halpern determined three ways in which this question can be answered. The first of these is to give psychometrically sound intelligent tests to large samples of men and women. Hence, the sex that obtains the higher average score is smarter. However, in her criticism of those tests, Halpern says that “...all of the major intelligence tests have been written so that there is no overall sex difference” (ibid, p.96). In a similar thought, Becker (1986) asserts that: “Sex differences in general intelligence are not found, but the tests were designed that way. Items showing sex differences were either eliminated or balanced by items showing a difference favoring the other sex” (p.367). Hence, there is a kind of agreement on the unreliability of these test results in the way that they are intended to show a balance in males` and females` general intellectual capacities. Yet, according to Halpern (2012), there are some researchers such as

Rushton and Jensen (2006), who maintain the idea that women are less intelligent than men and, who “...also maintain that Blacks are less intelligent than Whites and Whites are less intelligent than Chinese” (p.96). Such judgements can give hints about the basis and the reliability of their conclusions.

The second way to determine which sex is smarter on average might be by looking at the jobs or roles in society occupied by both men and women. However, this cannot be taken as a reliable source of evidence as this is more dependent on societal and cultural factors rather than cognitive abilities.

A third way suggested by Halpern (2012), and this one is of great importance as it is directly related to the topic of the present study, is to look at school achievement. In this regard, Halpern raise a pertinent question that she answers in the following way:

Which sex, on the average, gets better grades in school and is more likely to obtain advanced degrees? Numerous sources of data clearly show that women get better grades in school than men in every subject area, although the differences are not large. (p.93).

Yet, she later clarifies that “The data favor the intelligence of females when looking at course grades, and they favor males when looking at most high stakes exams such as the Advanced Placement tests and SATs” (ibid).

1.1.2.3.2. Perception

After reviewing some studies related to sex differences in terms of *audition*, *olfaction* (*the ability to detect faint smells*), *vision*, *taste*, and *time perception*, Halpern (2012) ended up with the following conclusion:

Reading through a long list of ways that males and females differ in perception can create a strong belief that women and men are living in separate worlds where sounds have different qualities, colors have different hues, and food has different tastes. But, in reality, all of the

sex differences in perception are quite small, even for the most reliable ones such as differences in the ability to smell faint scents, remember odors, or estimate time intervals. Even more importantly, these differences have no effects for the vast majority of people in their day-to-day lives... there are many more ways in which males and females are similar in perception than there are ways that they differ. (pp. 113-114)

Halpern explains that these data on sex differences in perception have been misused and wrongly interpreted by many people in the field of education. She argues that perception thresholds are measures of the minimum amount of stimulation needed for someone to perceive it, and for example, the fact that girls are often able to perceive light or sound stimuli at lower levels than boys does not mean that they live in perceptually different worlds. In this concern, Halpern (2012) points out that “The claims that boys and girls need different types of education based on sex differences in perception are based on faulty reasoning and, in some cases, fiction” (p. 114).

1.1.2.3.3. Memory

Males and females are not identical in terms of memorizing knowledge and procedures. Halpern (2004) confirms that “Compared with men, women have more rapid access to phonological, semantic, and episodic information in long-term memory” (p.136). In the same concern, she states that “males have the advantage on tests of verbal analogies, which may seem to be verbal but at a cognitive-process level involve mapping relationships in working memory” (ibid).

Likewise, Ullman et al. (2008), after exposing some studies on the differences in memory between males and females, ended up with the following conclusion:

In sum, previous evidence suggests the existence of both behavioural and neural sex differences in language. We have suggested that this pattern of data, as well as a range of

other evidence, can be explained by a female advantage at declarative memory, perhaps accompanied by a male advantage at procedural memory. (p.302)

1.1.2.3.4. Verbal Abilities

The term '*Verbal abilities*', just like the other cognitive abilities, is not a unitary concept as it applies to all the components of language usage, which are: word fluency, grammar, spelling, reading, writing, verbal analogies, vocabulary, oral comprehension as well as dealing with separate brain subsystems that are involved in generating language, comprehending language, using grammatical rules, and in producing and decoding speech sounds (Halpern, 2012).

Women, according to Halpern (2004, p.136) "...obtain higher scores on tests of verbal learning and the production and comprehension of complex prose". However, Hyde and McKinley (1997) have another opinion. After presenting the results of some studies, they concluded that:

...based upon the meta-analyses reported by Hyde and Linn (1988) and by Hedge and Nowell (1995), gender differences with regard to general verbal ability appear to be so small that they are now essentially non-existent. There are, however, many distinct types of verbal ability, and genuine differences may exist within particular areas such as speech production. Moreover, there are marked gender differences in disturbances of verbal functioning, where more boys exhibit difficulties than girls. (p. 34)

To support this, the results of the study by Hyde and Linn (1998) in which 165 studies reporting data about gender differences in verbal ability are presented here. Hyde and Linn say that:

Analyses of effect sizes for different measures of verbal ability showed almost all to be small in magnitude: for vocabulary, $d = 0.02$; for analogies, $d = -0.16$ (slight male superiority in performance); for reading comprehension, $d = 0.03$; for speech production, $d = 0.33$ (the largest effect size); for essay writing, $d = 0.09$; for anagrams, $d = 0.22$; and for tests of general verbal ability, $d = 0.20$. For the 1985 administration of the SAT - Verbal, $d = -0.11$, indicating superior male performance. Analysis of tests requiring different cognitive processes involved in verbal ability yielded no evidence of substantial gender differences in any aspect of processing. Similarly, an analysis by age indicated no striking changes in the magnitude of gender differences at different ages, countering Maccoby and Jacklin's conclusion that gender differences in verbal ability emerge around age 11. (pp. 2-3)

Halpern (2012) concludes her discussion about differences between males and females based on verbal abilities by making the link with the occupied roles and professions, which she considers as being a 'paradox'. She points that:

The research evidence from a variety of sources favors female superiority on verbal tasks including reading and speaking and is largest in the high and low ends of the distribution. Despite the finding that females score higher on at least some tests of verbal ability, the overwhelming majority of critically acclaimed writers are male. Other careers and prestigious professions that require advanced verbal abilities, careers like lawyer, politician, and journalist, are also predominantly male. (p. 123)

1.1.2.3.5. Visuospatial abilities

Visuospatial skills, or spatial skills that are visual in nature, are used extensively in domains such as engineering, architecture, chemistry, building...etc. Becker (1986, p. 368) argues that "The ability to manipulate two- and three-dimensional figures in your head (as an architect or mechanical engineer might have to do), could be strengthened by the kinds of toys

and play activities boys get into”. This, according to Becker’s words, implies that boys are better than girls in such abilities.

On the other hand, Hyde and McKinley (1997) argued, again, that, based on the findings of Linn and Petersen (1985) and Voyer, Voyer, and Bryden (1995), “...it makes no sense to make general statements concerning gender differences in spatial ability” (p. 37). Hyde and McKinley noted that both studies concluded that there are three main types of spatial ability: 1- *spatial perception*, which taps for example the participant's sense of horizontality or verticality; 2- *mental rotation*, which is measured by procedures that involve mentally rotating a three-dimensional object that is depicted in two dimensions to see whether it matches one of some other illustrations; 3- *spatial visualization*, is measured by procedures that involve visually locating a simple figure within a complex one. Richardson (1997) reports that, in all these tests, “Men tend to perform better than women do on some measures of spatial ability, but the magnitude of the gender difference varies markedly with the demands of each specific test” (p.131).

In addition to the three types of spatial ability previously stated, Halpern (2012) reported from some other studies, which dealt with two other types: *Spatiotemporal ability* involves judgments about and responses to dynamic visual displays, and the *Generation and maintenance of a spatial image*, which requires participants to generate an image (either from long-term or short-term memory) and then use the information in the image to perform a task. Halpern concludes her discussion by asserting that “Regardless of which definition is used, sex differences in spatial tasks are among the largest sex differences” (138).

1.1.2.3.6. Quantitative (Mathematical) abilities

In this particular and essential type of cognitive ability, the results of the numerous studies conducted to scrutinise the degree of gender differences varied in terms of focus. Some

studies such as Hyde et al. (1990) and Hedges and Nowell (1995) analysed gender differences with regard to performance in mathematics while other studies like Hyde et al. (1990) were more complex as they considered that problems in mathematics tap different cognitive levels, which are computation, concepts and problem-solving (Hyde & McKinley, 1997). Halpern (2012, p.147) gave some examples of tasks (used to assess quantitative ability) that belong to different mathematical areas which are: algebra, geometry, geometry (advanced and highly spatial), computation, word problem, calculus, mathematical proof, differential equations (in quantum mechanics).

Richardson (1997, p.143), in his conclusions from gender differences in cognition studies, reports on this point by saying that “There are essentially no consistent gender differences in measures of mathematical ability, with the sole exception that beginning in high school men tend to perform better than women do on tests of mathematical problem solving”. However, Halpern (2004), contrary to the belief that males outperform females in quantitative skills, surprisingly points that:

Girls have the advantage on quantitative tasks in early elementary school when math involves learning math facts and arithmetic calculations, showing rapid memory retrieval similar to that needed in language production and comprehension, and in later grades they perform better than males on algebra problems when the cognitive components of the solution strategy are similar to those of language processing. (p.136)

In the same concern, Halpern (2012) reported from different studies dealing with gender differences in mathematical abilities such as Hyde, Lindberg, Linn, Ellis, and Williams (2008); Guiso, Monte, Sapienza, and Zingales (2008); and Else-Quest, Hyde, and Linn, (2010). The conclusion was that “the average differences between females and males in math are small, but as the samples become more select, that is as ability levels increase, males outscore females” (p.148).

1.1.2.4. Psychological Differences

Discussion, as far as the psychology of both sexes is concerned, will turn around the degree of masculinity and femininity of a person. The newborn baby, as stated before, is categorised as being male or female according to apparent biological features or external genitalia and this classification of sexes is referred to as *a medico-legal category* (Meyer, 2010). The physical growth of the child is necessarily accompanied by psychological development and the construction of some features and personality traits that reveal the degree of concord between a person's sex and their gender.

According to Becker (1986), females are less confident and more anxious than males in problem-solving tasks. He also points out that males are more motivated by challenges while females are motivated greatly by social reinforcers rather than self-involvement. Becker summarises some of the main differences between boys and girls by saying:

Girls show more dependence, anxiety, fear, and compliance. They are more relationship-oriented and less achievement-oriented than boys. Women are more likely than men to end up in outpatient therapy for a variety of problems like phobias, depression, anorexia, and agoraphobia. (p. 369)

1.1.2.5. Sociological Differences

1.1.2.5.1. Behaviours

A concept that is strictly related to gender differences in terms of behaviours is that of *heteronormativity* or *heteronormative gender order* advanced by Holmes (2007), who explains it as follows:

An order based on the idea that there are two opposite sexes that are attracted to each other. The gender order demands that we categorize people as women or men. People usually try to imitate what are perceived as 'normal' femininity or masculinity and the complex

intersections between gender and sexuality are key to how this operates. Social rules about ‘normal’ gender and sexuality demand that you must be clear about who are boys and who are girls, so that boys and girls can grow up, fall in love with each other, and have more little boys and girls. (p.21)

Some differences between males and females can be depicted in behaviours such as aggressiveness, dominance and competitiveness. Men care more about dominance, and not surprisingly, they also like competition more than women do (Rhoads, 2004). Brannon (2017) made a comparison between men’s and women’s behaviours, from a different perspective. She considers that:

When men and women are put into situations without gender-related cues, their behavior tends to be quite similar. For example, a literature review (Frodi, Macaulay, & Thome, 1977) and meta-analyses (Archer, 2004; Eagly & Steffen, 1986; Knight, Guthrie, Page, & Fabes, 2002) of aggression have shown that women and men are similar in their experience of anger and willingness to behave aggressively in laboratory situations, but outside the laboratory, gender differences appear. One prominent difference is the type of aggression; men tend to choose direct confrontation whereas women are more likely to use an indirect strategy of causing harm. (p.511)

When it comes to activity level, males are more active, and curious and show exploratory behaviour, especially in the period of childhood. This applies to impulsivity too as observational studies show that males are more impulsive (Becker, 1986).

1.1.2.5.2. Language Use

Haas (1979) considers that male and female speech tend to differ in terms of four aspects: form, topic, content, and use. In her article, Haas reviewed many studies that compared male and female language and came up with the conclusion of the empirical evidence, which implies that:

Men may be more loquacious and directive; they use more nonstandard forms, talk more about sports, money, and business, and more frequently refer to time, space, quantity, destructive action, perceptual attributes, physical movements, and objects. Women are often more supportive, polite, and expressive, talk more about home and family, and use more words implying feeling, evaluation, interpretation, and psychological state. (p. 616)

Sunderland (2006) summarised the aspects of language use in which the main differences between men and women are depicted in the following passage:

The topic of women's and men's speech has been of particular interest to sociolinguists. Issues include gender-differential tendencies in style-shifting (for example, between formal and casual speech), use of prestige and stigmatised variants, linguistic conservatism, who leads language change (see below) and the positive and negative evaluation of such change. (p.7)

She reported from Trudgill (1972) that women tended to use the prestige form more than men and they also tended to over-report their pronunciation, that is when asked about their pronunciation, they produced more 'prestigious' sounds than they actually did. Concerning linguistic conservatism, Sunderland referred to the works of Labov (1962, 1966, 1972, and 1990). In his study, Labov's (1966) found that "...women in New York tended to style-shift far more than did his male informants and that they tended to be less conservative linguistically" (ibid.). According to Sunderland, a contrasting result was found in Labov (1962, 1972), who noted that men were leading the change. By reporting on the study of Labov (1990), Sunderland (2006) noted that "...whereas men tend to use more non-standard stable forms than women, when it comes to linguistic change, women will innovate, in part by favouring new prestige forms" (p.8).

1.1.2.5.3. In-home/Outside Home Roles

Gender stereotypes play a great role when it comes to assigning roles to men and women in society. The stereotypical image held in most societies is that women are housekeepers, child carers and responsible for the well-being of their husbands, especially in terms of the emotional and mental aspects. In this respect, Rhoads (2004) says:

The biggest issues in contemporary marriages surround women's work, in and out of the home. Women like to care for children much more than men do, and they also like housework more than men do. When vocational counsellors use the Strong Interest Inventory to help guide clients to appropriate occupations, "Some of the largest sex differences are found on such tasks as cooking, sewing and 'home economics.'" There is reason to think that such differences have a biological basis. Preteen and teenage girls exposed to unusually high levels of testosterone *in utero* are more interested than other girls in a career other than homemaking. Correspondingly, women who have Turner's syndrome and thus produce no testosterone typically have unusually high interests in housekeeping.

Similar to in-home roles, choices of outside-home roles and activities may reflect genetic or other biological bases and are greatly influenced by gender stereotypes, which implies that these choices are not made on a voluntary and free basis (Brannon, 2017). Giving the situation in the United States as an example, Brannon (2017) clarifies that:

Although men and women have similar mathematical ability and now have similar preparation in terms of courses completed in high school (Spelke, 2005) and an increasing number of college degrees in math (U.S. Census Bureau, 2015), young men are more likely than women to pursue careers that rely on math. Even women who complete the courses required for a good math background do not choose science and engineering careers as often as men do (Ayalon, 2003; Ceci, Williams, & Barnett, 2009). (p.512)

Dealing with the same point of choices, Brannon considers that:

The movement of women into the paid workforce has increased their options in some ways but not in others. Women have more choices than men do concerning employment or homemaking—women may be homemakers or employed, or alternate the two, but men are still expected to be employed. (ibid).

1.1.3. Gender Similarities

There is very little literature dealing with gender similarities than that addressing gender differences. Richardson (1997) comments on this by saying that “research studies that reveal no significant difference between males and females might be less likely to be accepted for publication” (p. 9).

Hyde (2005, p.581), advanced a different view to that of gender differences that captivates the media and the general public as a whole. She discussed the so-called *Gender Similarities Hypothesis*, which “...holds that males and females are similar on most, but not all, psychological variables. That is, men and women, as well as boys and girls, are more alike than they are different”. In terms of cognitive abilities, for example, Halpern (2012) explains that she has:

...focused on differences because the logic of hypothesis testing only allows conclusions about differences. Despite this limitation, it is important to note that the number of areas in which sex differences are even moderate in size is small. Males and females are overwhelmingly alike in their cognitive abilities. (p.155)

Section Two: The Place of Gender within the Algerian Curriculum

1.2.1. Definition of Curriculum

Allen (1984, p.61) defines *curriculum* as "...a very general concept which involves consideration of the whole complex of philosophical, social and administrative factors which contribute to the planning of an educational programme". According to Nunan (1988), "*Curriculum* is concerned with the planning, implementation, evaluation, management, and administration of education programmes" (p.8).

Richards and Schmidt (2010) define a *curriculum* as "An overall plan for a course or programme" (p.151). They clarify that such a programme usually states:

- a. the educational purpose of the programme, in terms of aims or goals;
- b. the content of the programme and the sequence in which it will be taught (also known as the syllabus);
- c. the teaching procedures and learning activities that will be employed (i.e. methodology);
- d. the means used to assess student learning (i.e. assessment and testing);
- e. the means used to assess whether the programme has achieved its goals (i.e. evaluation).

(ibid)

1.2.2. Language Curriculum Development

In his discussion of curriculum development in second language teaching, Richards (1990) states that:

Language curriculum development, like other areas of curriculum activity, is concerned with principles and procedures for the planning, delivery, management, and assessment of teaching and learning. Curriculum development processes in language teaching comprise needs analysis, goal setting, syllabus design, methodology, and testing and evaluation. (p.1)

Thus, Richards, in the last statement, lists five steps in curriculum development:

- **Needs analysis:** a set of procedures used to generate a considerable amount of data about the context of the language programme, the learners, the teachers, and the administrative factors affecting the programme.
- **Goal setting:** general statements of the intended outcomes of the language programme based on the curriculum planners' and/or policymakers' beliefs about the desired aims of this programme.
- **Syllabus design:** put in simple words, this refers to the specification of the instructional content of the programme.
- **Methodology:** taken as a separate step from syllabus design, methodology refers to the activities, tasks, and learning experiences chosen and employed by the teacher to achieve learning.
- **Testing and evaluation:** in the process of curriculum development, testing can often be a component of needs analysis and evaluation as well. The focus of this latter is to determine whether the programme is effective i.e. to check the extent to which its aims have been achieved.

Richards (2001) considers that within the process of curriculum development, teachers and language teaching professionals should be placed at the centre of the planning and decision-making processes. Many people with different roles take part in these processes, resulting in products that, according to Richards (2001, p.42), "...are easy to identify and analyse because they exist in the form of policy documents, syllabuses, tests, teaching materials, teaching programmes, textbooks, and teaching and learning acts". The following table from Johnson (1989, p.3) illustrates these decision-making roles and their products:

Table 1.1

Stages, decision-making roles and products in curriculum development.

Note. Adopted from Richards (2001, p.42)

Developmental stages	Decision-making roles	Products
1- Curriculum planning	Policymakers	Policy document
2- Specification: ends means	Needs analyst	Syllabus
	Methodologists	
3- Programme implementation	Material writers	Teaching materials
	Teacher trainers	Teacher-training programme
4- Classroom implementation	Teacher	Teaching acts
	Learner	Learning acts

The greatest impact of needs analysis has been noticed in the field of English for Specific Purposes. This does not reject its important role in planning general English courses. In this section, the interest is on the goals of the programme as it focuses on the analysis of the aims of the Algerian curriculum as a whole (i.e. the policy of the educational system) as well as the aims of the English language curriculum in particular. More focus is put on the last three steps of curriculum development, i.e. syllabus design, methodology and testing and evaluation, in Chapter Two.

1.2.3. The Algerian Curriculum: A Reflection on the Educational Policy

The ministry of national education (2008) published the *National Education Orientation Law*, which is the latest document orienting the educational policy of the country. It acts both as a guideline and a general framework for the whole Algerian educational system. This document sets the principles and the primary aims of the new educational policy in Algeria, which are summarised as follows:

- Enhancing the school's role in proving the Algerian identity and consolidating the unity of the Algerian people. This is to be done by promoting the values related to Islam being a religion, a culture and a civilisation, Arabic being a language, culture and civilisation and also Tamazight as a language, culture and heritage;
- Ensuring training on citizenship;
- The integration of the Algerian school in the global advancement movement. Within this aim, lies the importance given to learning foreign languages as a necessity in a world witnessing economic, technological, scientific and cultural transformations;
- Emphasising and enhancing the principle of democracy;
- Valuating and promoting human resources.

The National Education Orientation Law (Ministry of National Education, 2008) consists of seven chapters organised as follows:

- **Chapter one:** The principles of the Algerian school;
- **Chapter two:** The educational community;
- **Chapter three:** School organisation;
- **Chapter four:** Adult education;
- **Chapter five:** The personnel. The focus here is to ensure the best training for educators to achieve not only *the same education for all* but also *the best education for everyone*;
- **Chapter six:** The educational institutions, support structures and activities, and advisory bodies;
- **Chapter seven:** Transitional and final provisions.

Based on the National Education Orientation Law (Ministry of National Education, 2008), 'the national curriculum committee', which was installed in November 2002, and under the supervision and authority of the Ministry of National Education, published another

document entitled *The General Reference for the Curricula (March 2009)*. This reference document, according to its authors, acts as a reminder about the ends of the educational reform that are closely related to the aims of the Algerian school. Those ends, which are also dealt with in the 2008 Law referred to before, were determined by the decisions of the ministers' council meeting on April 30th, 2002, and have their backgrounds in the recommendations of the National Committee for the Educational System Reform set in May 2000. This general reference:

تحدد...الخطوط العريضة القوية التي ينبغي أن يحملها التغيير المستهدف بالإصلاح، مع اقتراح أساليب جديدة للتفكير، ومقاربات جديدة للتنفيذ، وتنبثق عنها كل الوثائق المنجزة انطلاقاً من التوجيهات الدقيقة التي تقدمها، وكذا الطرق التقنية لتصوير وإنجاز المناهج الجديدة.

[...sets the basic outlines that the reform's targeted change should carry, proposing new ways of thinking and new approaches for the implementation; and represents also the source of all the achieved documents due to the precise guidance it provides, as well as the technical methods for conceptualizing and implementing the new curricula.] (p.62)

The two aforementioned policy documents are considered the cornerstone for the curricula of all the school subjects in primary, middle and secondary education in Algeria. The general frameworks of both the primary school and middle school curricula, for example, refer to the National Education Orientation Law (Ministry of National Education, 2008) as it determines the ends targeted by the educational system as well as the principal roles of the Algerian school. In addition to these basic elements, these general frameworks review the founding principles of the curricula, which can be grouped, according to their authors, into three main fields: Ethical, epistemological (philosophical), and methodological and pedagogical. These, they argue, are to be taken into consideration in addition to *inclusivity*, *harmony*, *feasibility*, *readability*, *relevance*, and *evaluability* which are the basic principles of any curriculum design process.

Moreover, these general frameworks present the components of the curricula. These components are listed below:

- Exit profiles;
- Resource blueprint for competencies' building / Cognitive Resource Matrix;
- Cross-sectional values, competencies, and common themes;
- Annual programme schedule;
- Learning;
- Assessment.

1.2.4. The English Language Curriculum in Algeria

The teaching-learning of the English language in the Algerian educational system starts from the first year in middle school. The latest English language curriculum for middle school was released by the ministry of national education in 2016 in a document entitled: *The Educational Programs for the First Stage of Middle School Education*. The curriculum authors state that:

Rewriting the curriculum for teaching - learning English in Algeria meets with the coherent legislative framework regulating our educational system. The Orientation Law on National Education No. 08_04 23 (January 2008) having set four missions to the school: instruction, education and qualification, the new English curriculum fully integrates these missions and subscribes to the official educational guidelines. The four years of the middle school are important foundation years for the intellectual, emotional, physical and social development of the child. These years will help him* (*he/him are unmarked forms used here to refer to both male and female) to gradually acquire competencies at all levels of school education and to continue learning even after leaving school. (p.90)

They consider that the teaching of English at the middle school must meet the following objectives:

- To help our society to live in harmony with modernity by providing the learner with linguistic tools essential for efficient communication;
- To promote national and universal values;
- To develop critical thinking, tolerance and openness to the world;
- To contribute to the shaping of a good citizen, aware of the changes and challenges of today and tomorrow;
- To give every learner the opportunity to have access to science, technology and world culture while avoiding the dangers of acculturation. (ibid)

The curriculum authors add that:

Teaching-learning English allows the learners to interact with peers and other English-speaking persons, using their competencies of interaction, interpretation and production that they can develop through the oral or written mode in and out of school. Like other disciplines, teaching-learning English helps the acquisition of values as well as cross-curricular competencies. These competencies are four in number: intellectual, methodological, communicative, personal and social. (ibid)

In the secondary school phase, the newest versions of the curricula were published in January 2005, January 2006 and January 2007 for the first, second and third-year secondary school levels respectively. In addition to the general objectives for the teaching of English in each of these secondary school levels, the set of these curricula presents the general objectives of teaching English in the secondary school as well as the targeted learner's exit profile, which they refer to as the *Integration Terminal Objective*.

The English language curriculum for the first year of secondary school (Ministry of National Education, 2005), for example, determines the general aims (ends) of the teaching-learning of English in this cycle as well as its general objectives. The curriculum authors summarise these aims by saying:

The teaching of English must participate in the development of the learner in all its dimensions. In addition to the national values, it advocates openness to the world, respect for oneself and others as well as tolerance.

Intervening as a second foreign language and covering seven years of study (four in the middle cycle and three in the secondary cycle), the teaching of the English language aims to establish and develop communicative, linguistic and, cultural and methodological competencies which will allow the learner to deal with situations of oral and / or written communication taking into account his future needs and those of the society in which he evolves. Thus, the teaching/learning of the English language will allow the learner to communicate, exchange, and immerse themselves in the culture conveyed by this language and to use it as a cultural, scientific and technical tool. (p.4)

They also classified the general objectives of the teaching-learning of English in secondary school into four main categories:

1- Linguistic (school) objectives

- To provide the learner with the necessary tools to pursue higher education in English (teaching license, interpreting, etc.)
- To promote the development of basic skills to understand and communicate.

2- Methodological objectives

- To promote autonomous learning strategies in learners to enable them to deepen and broaden their knowledge.
- To strengthen the learner's mental and intellectual skills such as analysis, synthesis, and evaluation through relevant activities.

3- Cultural objectives

- To encourage the learner to explore different cultural aspects of other linguistic communities and thus to better understand and apprehend their own culture.
- Fostering positive attitudes towards others.

4- Socio-professional objectives

- To allow the learner to take advantage of the various documents in English (literary, technical, scientific, economic documents and cultural productions) that he/she might encounter and that he/she would have to consult in his/her professional and/or university life.

The learner's exit profile from the secondary school cycle is determined in the curricula of three school levels. The third-year secondary school English language curriculum (Ministry of National Education, 2007), for example, states that:

The Terminal Integration Objective that the student learning English must achieve at the end of his/her three years in the secondary school cycle is the following:

In a communication situation, and on the basis of an oral or written support, the student must produce a written message of about twenty lines, in a chosen type of written discourse (descriptive, narrative, argumentative, expository, injunctive), correctly and legibly. (p.9)

1.2.5. Learners' Gender within the Algerian Curriculum

The National Education Orientation Law (Ministry of Education, 2008, p.39) states clearly, within its framework, that one of the main aims of the Algerian educational system to be fulfilled in the mid and far terms are:

- العمل على تجانس وتحسين معايير التمدرس داخل الولاية الواحدة و بين ولايات القطر الوطني والتقليص المعتبر للتسرب المدرسي خلال التمدرس بشكل يمكننا إلى غاية 2015 من توصيل 90% من دفعة السنة الأولى ابتدائي إلى السنة الرابعة متوسط.

- ضمان تساوي نسبة التمدرس بين البنات و البنين في مختلف مستويات التعليم.

Working to harmonize and improve schooling standards within the same state and between the states of the country and to reduce the significant dropout during schooling in a way that we will be able until 2015 to get 90% of the first-year primary school classes to the fourth year. Ensuring an equal schooling rate between girls and boys at all levels of education.

This reflects the awareness of the Algerian educational policymakers of the danger of delinquency among learners and the importance of giving equal chances of success to ALL Algerian pupils as well as the great importance given to eliminating the gender gap in the rates of males and females schooled in compulsory education.

These intentions of the educational policy makers manifested in the National Education Orientation Law had to be put into practice in the field. Hence, measures to fight delinquency among pupils from both genders had to be taken in addition to alternative instructions, materials and practices to guarantee equal chances of success for pupils from different provinces, different backgrounds, and from both sexes.

The learners' gender is depicted in another place in the policy documents designed by the ministry of national education. *Article 12* of the National Education Orientation Law (Ministry of National Education, 2008) states that “التعليم اجباري لجميع الفتيات والفتيان البالغين من العمر (6) سنوات الى ست عشر (16) سنة كاملة [Education is obligatory for all girls and boys aged between 6 and 16 years old]” (p. 65). In this article, the age range of obligatory schooling is determined and the girls are mentioned in the first place. This can reveal the great importance given by the state to the schooling of female learners due to some societal and cultural factors that can lead parents to either inhibit their daughters, more than sons, from going to school or oblige them to leave school at an early age.

Conclusion

Boys and girls, according to the various pieces of research and meta-analyses studies referred to in this chapter, tend to show more similarities than differences, especially in terms of the cognitive skills of intelligence, perception, verbal abilities and quantitative abilities. More noticeable differences are depicted when they were compared based on visuospatial abilities. Not surprisingly, boys and girls differ in terms of biological, physiological and psychological features. Gender stereotypes also play an important role in enhancing social differences

between males and females in terms of expected behaviours, language use, as well as men's and women's roles inside and outside the home. These findings about gender similarities and differences are of paramount importance for educational policy-makers and curriculum designers as they provide them with a platform to deal with the issue of the learners' gender within the process of setting up the educational policy and the curriculum development process as a whole with its constituents such as the curriculum, the syllabus, the instructional materials, the tests...etc. These are to be dealt with in detail in the next chapter.

Chapter Two

Syllabus Design, Teaching/Learning Materials and Assessment & Evaluation: Basic Components of the Curriculum

Introduction	37
2.1. Syllabus Design	37
2.1.1. Definition of a ‘Syllabus’	37
2.1.2. Syllabus Design vs. Methodology	38
2.1.3. Syllabus Design and Teaching Materials	39
2.1.3.1. Defining ‘Materials’	39
2.1.3.2. The Framework of Materials and Methods in ELT	40
2.2. Assessment/Evaluation	57
2.2.1. Types of Assessment	58
2.2.1.1. Informal/Formal Assessment.....	59
2.2.1.2. Formative/Summative Assessment	59
2.2.1.3. Norm-referenced/Criterion-referenced Assessment	60
2.2.1.4. Classroom-based/Large-scale Assessment	62
2.2.2. Test Types	62
2.2.2.1. Placement Test	62
2.2.2.2. Proficiency Test.....	63
2.2.2.3. Diagnostic Test.....	63
2.2.2.4. Achievement Test.....	64
2.2.3. Framework of Language Test Task’s Characteristics	65
2.2.4. Matching Tests to Decision Purposes.....	67
2.2.5. Cardinal Criteria for ‘Testing a Test’	69
2.2.5.1. Validity	69
2.2.5.2. Reliability	72

2.2.5.3.	Practicality	73
2.2.5.4.	Equivalency	74
2.2.5.5.	Washback	74
Conclusion.....		75

Introduction

The educational process is one in which a bi-directional relationship exists between its main constituents: The use of a particular type of *instruction* (through teaching based on some curricula) to achieve or arrive at a certain set of *standards* (either a language proficiency or academic content) that can be tested through *assessment* (whether classroom-based or larger scale) (Mihai, 2010). Moreover, Nunan (1988) explains that one of the different perspectives from which a curriculum can be looked at is that of planning. He assumes that decision-making is done “...in relation to identifying learners’ needs and purposes; establishing goals and objectives; selecting and grading content; organizing appropriate learning arrangements and learner groupings; selecting, adapting, or developing appropriate materials, learning tasks, and assessment and evaluation tools” (ibid, p.4).

Hence, in this chapter, these elements will be discussed by referring to syllabus design and the different types of syllabi, being a basic component and an application of the curriculum, approaches to syllabus and materials design, and materials’ evaluation as a key element in the sake of making the academic content adequate to the curriculum, the teaching/learning context and also to learners’ profiles. The last part of this chapter deals with assessment and evaluation in general and testing in particular.

2.1. Syllabus Design

2.1.1. Definition of a ‘Syllabus’

Breen (1984, p.47) defines a syllabus as “a plan of what is to be achieved through our teaching and our students' learning”. A similar definition is found in McDonough et al. (2013, p.11), who say that “The ‘syllabus’ can be seen...as the overall organizing principle for what is to be taught and learned”. Likewise, Nunan (1999, p.73) points out that “A syllabus consists of lists of content to be taught through a course of study. Key tasks for the syllabus designer are the selection of the items and their sequencing and integration”. This role of the syllabus

designer was the topic of conflict between scholars who have different views about the nature of syllabus design and methodology, especially with the rise of communicative language teaching.

2.1.2. Syllabus Design vs. Methodology

To clarify the nature of the relationship between syllabus design and curriculum development, Richards (2001, p.2) says that “Syllabus design is one aspect of curriculum development but is not identical to it”. Nunan (2015) discussed the components of curriculum development by saying:

There are three subcomponents to curriculum development: syllabus design, methodology, and evaluation. All of these components should be in harmony with one another: methodology should be tailored to the syllabus, and evaluation/assessment should be focused on what has been taught. (In too many educational systems, what is taught is determined by what is to be assessed. (p.5)

Two different views about syllabus design are found in literature: one is narrow and the other is broad. Nunan (1988) points that:

The narrow view draws a clear distinction between syllabus design and methodology. Syllabus design is seen as being concerned essentially with the selection and grading of content, while methodology is concerned with the selection of learning tasks and activities. Those who adopt a broader view question this strict separation, arguing that with the advent of communicative language teaching the distinction between content and tasks is difficult to sustain. (p.5)

Widdowson (1984, p.25), for example, seems to support the narrow view when discussing the dual relationship between policy and methodology, which he thinks makes the syllabus a source of conflict by saying “The design of a syllabus and its implementation by means of methodology can never be solely a pedagogic matter”. Yet, Nunan (1999, p.73) argues that

“...With the emergence of CLT has come a group of curriculum specialists who take a broader view, and who question the sustainability of this strict separation”.

2.1.3. Syllabus Design and Teaching Materials

Brown (1995, p.151) believes that “Closely related to syllabus design is the question of deciding what kind of organisational framework to adopt for developing materials”. To start with, a definition *materials* with a reference to *textbooks* is provided and then followed by a discussion of the framework of materials and methods in English language teaching (ELT).

2.1.3.1. Defining ‘Materials’

Brown (1995, p.139) defines materials as “any systematic description of the techniques and exercises to be used in classroom teaching”. Another very practical definition of *materials* is found in Tomlinson (2011):

Anything which is used to help language learners to learn. Materials can be in the form, for example, of a textbook, a workbook, a cassette, a CD-ROM, a video, a photocopied handout, a newspaper, a paragraph written on a whiteboard: anything which presents or informs about the language being learned. (pp. xiii-xiv)

For our research purposes, when we refer to materials, we are basically thinking about textbooks that are organised and sequenced around a particular syllabus. This view is shared by McGrath (2002) saying that the focus in his book is on *text* materials. He makes the point that “In many situations, the expectation is that teaching is based on a single textbook, although other materials may be used at the teacher’s discretion. The term ‘Coursebook’ will be used to refer to a textbook on which a course is based” (ibid, p.7).

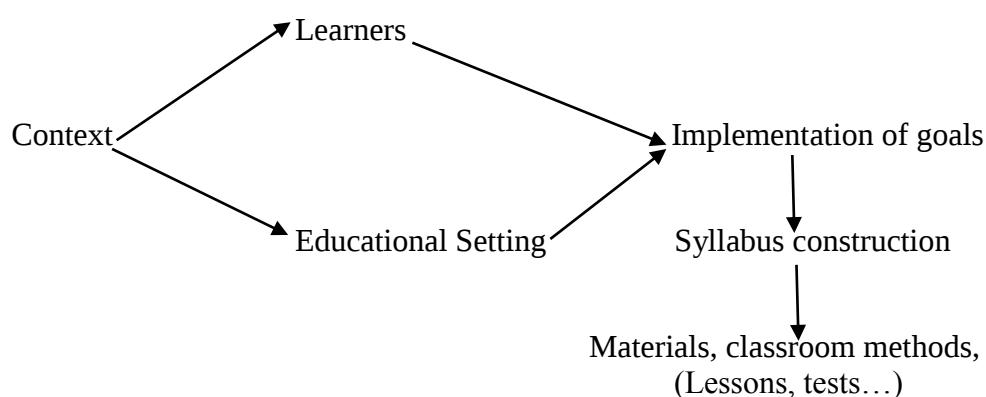
2.1.3.2. The Framework of Materials and Methods in ELT

McDonough et al. (2013) developed a framework for English language teaching that they summarised in the following figure:

Figure 2.1.

The framework of language teaching

Note. Adopted from McDonough et al. (2013, p.6)



In this framework, they consider that materials and methods cannot be seen in isolation. Yet, they are part of a broader professional context that is illustrated in figure 2.1., which represents the different stages of planning an English language programme. To explain these stages McDonough et al. (2013) state:

Whether goals are stated in terms of a national language policy, or in the more specific environment of, say, a particular school or college, the possibilities for actually implementing them will be directly related both to the learners themselves – their needs, characteristics and so on – and to the whole educational setting in which the teaching is to take place...The statement of goals, then, related to the learners and conditioned by the setting, leads to the selection of an appropriate type of syllabus content and specification. The broad syllabus outline will in turn have direct implications for the more detailed design and selection of materials and tests, the planning of individual lessons and the management of the classroom itself. (pp. 5-6)

In what follows, the most important contextual factors, and the available syllabus types are presented.

2.1.3.2.1. The Context

This, as the previous figure illustrates, is subdivided into two main categories: Learner characteristics and the educational setting. These are summarised in Table 2.1.:

Table 2.1.

Contextual Factors Involved in Planning a Language Programme

Learner Characteristics	Educational Setting
- Age - Interests - Level of proficiency - Aptitude - Mother tongue - Academic and educational level - Attitudes to learning, to teachers, to the institution, to the target language itself and to its speakers. - Motivation - Reasons for learning - Preferred learning styles - Personality	- The role of English in the country - The role of English in the school, and its place in the curriculum. - The teachers - Management and administration - Resources available - Support personnel - Size of classes - Time available for the programme - Physical environment - The socio-cultural environment - The types of tests used - Procedures (if any) for monitoring and evaluating the language teaching programme itself.

2.1.3.2.2. The Syllabus

McDonough et al. (2013, p.12), in their discussion of the different types of syllabi available as options for syllabus designers, clarify that:

Although syllabuses typically are written and published documents, their circulation is often restricted to the particular situation for which they have been drawn up. Therefore, one of the simplest ways of surveying the types of syllabus available is to examine the contents pages of published English language teaching textbooks, because they reveal the underlying principles and assumptions on which the writers have based their material.

Six syllabus types are, then, presented in this section: *structural/grammatical*, *situational*, *notional/functional*, *topical/content-based*, *skill-based* and *task-based*. Here, each of these six types is theoretically treated as if it occurs purely and in a distinct way from the other types. However, as Krahne (1987, p.8) argues, “In practice, of course, these different types rarely occur independently of each other. Almost all actual language teaching syllabi are combinations of two or more of the types defined here”. This stands as an argument for dealing with the *multi-syllabus syllabus*.

2.1.3.2.2.1. Structural/ Grammatical Syllabus

This type was the most commonly used and it is still widely referred to. It is organised based on a list of grammatical structures. According to Nunan (1988), the input in a grammatical syllabus:

...is selected and graded according to grammatical notions of simplicity and complexity...grammatical complexity does not necessarily equate with learning difficulty. In other words, what is grammatically complex will not necessarily be that which is difficult to learn, and that which is grammatically simple will not necessarily be that which is easy to learn. (p.28)

A grammatical syllabus designer assumes, then, that language consists of a finite set of rules, which can be learnt one by one and then combined to make meaning.

Today, few courses are planned solely based on grammatical criteria since traditional grammar-based syllabuses received much criticism. Richards (2001, p.153) summarized some of this latter in the following points:

- They represent only a partial dimension of language proficiency.
- They do not reflect the acquisition sequences seen in naturalistic second language acquisition.
- They focus on the sentence rather than on longer units of discourse.
- They focus on form rather than meaning.
- They do not address communicative skills.

However, no language course can neglect the grammatical component as part of its content regarding its paramount importance in language learning.

2.1.3.2.2.Situational Syllabus

This one presents the learners with a set of everyday situations or ‘settings’ (McDonough et al. 2013). Wilkins (1976) considers that taking into account *situational needs*, and therefore constructing *a situational syllabus*, is the most common to replace a grammatical syllabus, which he suggested to be the framework for most foreign language teaching. He believes that *a situational syllabus* “...is the only other kind of syllabus that is at all widely used as a basis for the construction of teaching materials” (ibid, p.16). The general assumption underlying the adoption of this syllabus type is that language is used in a social context and cannot be used in isolation. Hence, drawing on learners' needs, the situations that they need the language for are defined and on that basis, the language that is needed for those situations is taught. Units, then, will have situational rather than grammatical labels. Examples of those

labels include: At the airport, at the restaurant, booking a room at a hotel...etc. However, Wilkins criticises the situational syllabus by pointing that:

If we were to attempt to use a situational syllabus for any learner whose needs could not be identified in these situational terms, including the general language learner, we would fail to provide him with the means to handle significant language means. Useful as a situational syllabus may be in certain circumstances, therefore, it does not offer a general solution to problems of syllabus design. (ibid, p.18)

2.1.3.2.2.3.Functional / Notional Syllabus

In his distinction between the two terms of *functions* and *notions*, emphasising that they are confusing for many teachers on the first use, Nunan (1988, p. 35) argues that:

“In general, functions may be described as the communicative purposes for which we use language, while notions are the conceptual meanings (objects, entities, states of affairs, logical relationships, and so on) expressed through language”. As such, Krahne (1987, p.9) defines a functional/notional syllabus as one in which “the content of the language teaching is a collection of the functions that are performed when language is used, or of the notions that language is used to express”. Examples of these functions include: *identifying, agreeing, greeting, and requesting* while some examples of notions are: *cause, colour, duration, direction and time*.

2.1.3.2.2.4.Topical/Content-based Syllabus

Richards (2001, p.157) defines such a syllabus as “One that is organised around themes, topics or other units of content. With a topical syllabus, content rather than grammar, functions or situations is the starting point in syllabus design”. The point of departure for this syllabus type is its experiential content that is derived from a well-defined subject area that might be other subjects of the school curriculum like science, or any specialist field such as medicine, engineering, accounting...etc. (Nunan, 1988). Nunan, in his further discussion, gave as an

example the teaching of adult ESL classes in Australia most of which are characterized as being content-oriented. He argues that such syllabuses:

...take as their point of departure the skills and knowledge which syllabus planners and teachers feel is important for new arrivals. Units of work thus appear with labels such as 'health', 'education', and 'social services'. While the relevance of this content might seem obvious, many learners are confused by content-oriented courses, thinking they have strayed into a settlement rather than a language programme. In such cases, it is important for teachers to negotiate with the learners and demonstrate the relationship between language and content. (ibid, p.49)

Krahnke (1987) gave another sophisticated definition of content-based syllabi by specifying that:

Any content-based syllabus is by definition identical to the syllabus of a content course at any level in science, social studies, or any other school subject. Extensive reading of literature or other content material in a target language can also be seen as a type of content-based learning. A content syllabus might be supplemented with traditional, form-focused, language intensive work on, for example, vocabulary development, spelling, specific and intensive writing activities, and so on. (p.69)

Brown (1995, p.9) compared this syllabus type with the situational one and defined it by saying that "Topical syllabuses are similar to situational syllabuses. However, they are organised by topics or themes, rather than situations". He argues that these syllabuses are used to organise language texts that the textbook authors select and sequence based on the importance they perceive these topics or themes have on the students' lives. In fact, and contrary to what Brown stated, this type of syllabuses is not only used to organise reading texts but in many textbooks nowadays, they are found to be the basis for organising whole units with their reading texts, communicative activities and tasks, writing assignments...etc. Such a topic-based

organisation is found in the *New Headway Pre-Intermediate Student Coursebook*, where the first six units, for example, have the following headings:

- 1- You and Me
- 2- A good Job!
- 3- Work Hard, Play Hard!
- 4- Somewhere to Live.
- 5- Super Me!
- 6- Life's Ups and Downs

2.1.3.2.2.5.Skill-based Syllabus

This type focuses on the four language skills, and its main concern is what the learners can do with the language as speakers, listeners, readers and writers (McDonough et al., 2013). Examples of these skills include *reading skills* such as *skimming* and *scanning*, *listening skills* e.g. listening to foreign news reports for specific information, *writing skills* like writing memos, short reports or describing processes, and *speaking skills* such as ordering food from a menu, asking for and giving personal information...etc. *Skill-based instruction* is often equated with *competency-based instruction*, which refers to what a learner can *do* as a result of instruction (Krahnke, 1987). One of the settings in which such a syllabus can be encountered is while preparing a second or foreign language student to undertake a higher education course. Here, there is a need for training them on some skills such as *note-taking*, *skimming and scanning*, *oral presentation and public speaking*, *essay writing*...etc.

2.1.3.2.2.6.Task-based Syllabus

As its name indicates, this syllabus type uses *tasks* as its point of departure. A *task* has been defined by many researchers and scholars working in the field of language teaching, education, psychology and many other fields. Prior to citing any definition given to the term, it is worth mentioning that a clear distinction between *pedagogical tasks* and *real-world tasks*

(*target tasks*) has been done by scholars. Nunan (2004, p.1), for example, says that “target tasks, as the name implies, refer to uses of language in the world beyond the classroom; pedagogical tasks are those that occur in the classroom”. Since most language learning and teaching take place inside the classroom, the debate is more about the definition of pedagogical tasks and their characteristics. In this concern, Nunan (1989) makes a distinction between *communicative* and *non-communicative tasks*. Although he admits that it is not always easy to make such a hard and fast distinction, he considers a *communicative task* to be:

...a piece of classroom work which involves learners in comprehending, manipulating, producing or interacting in the foreign language while their attention is principally focused on meaning rather than form. The task should also have a sense of completeness, being able to stand alone as a communicative act in its own right. (ibid, p.10)

Littlejohn (2011), on the other hand, explains that restricting the scope of a ‘task’ to being a meaning-focused activity will be too narrow as a definition in attempting to analyse any set of materials. He opted for a broader one saying that a ‘task’ refers to:

...any proposal contained within the materials for action to be undertaken by the learners, which has the direct aim of bringing about the learning of the foreign language. Such a wide definition as the one above has the virtue of recognising that there may be many different routes to classroom language learning, from large-scale ‘whole tasks’ to short ‘gap fill’ exercises... (p. 188).

Another central point to refer to here in the process of dealing with *task-based syllabuses* is *Task-Based Learning and Teaching (TBLT)*. This term, according to Samuda & Bygate (2008, p.58), is used “to refer to contexts where tasks are the central unit of instruction: they ‘drive’ classroom activity, they define curriculum and syllabuses and they determine modes of assessment”. According to Brown (1995):

Authors who favour task-based syllabuses organise materials around different types of tasks that the students might be required to perform in the language. Such tasks might include *reading job ads, making appointments, writing a résumé, filling out a job application, being interviewed, solving a problem*, and so on. An author's selection of the tasks to be included in a task-based syllabus is typically based on their perceived usefulness to students. (p.11)

Some classifications of tasks and activity types, which engage the learners in practising the target language inside the classroom, are discussed in the literature. This is because of the many scholars who have written about TBLT. In Nunan (2004), some of the classifications, which are based on communicative language use, such as Prabhu (1987), Pattison (1987), and Richards (2001), are referred to. One more typology of tasks suggested by Ellis cited in Tomlinson (2011) is also referred to as follows:

- **Communicational Teaching Project Typology**

Prabhu's book published in 1987 "aims to present a particular view of second language pedagogy, together with an account of a five-year project of exploratory teaching which helped to articulate that view and to develop procedures of teaching consistent with it" (p.1). The project is frequently referred to in the literature as the Bangalore Project (e.g. Nunan, 2004); however, Prabhu (1987) made it clear that "the project team itself used the name Communicational Teaching Project" (p.1). A task, referred to by Prabhu as a 'Meaning-focused activity', can broadly be divided into *information gap, reasoning gap* and *opinion gap*. He explained this as follows (Prabhu, 1987, pp.46-47):

1. Information-gap activity, which involves a transfer of given information from one person to another – or from one form to another, or from one place to another – generally calling for the decoding or encoding of information from or into language. One example is pair work in which each member of the pair has a part of the total information (for example an incomplete picture) and attempts to convey it verbally to the other. Another example is completing a tabular

representation with information available in a given piece of text. The activity often involves the selection of relevant information as well, and learners may have to meet the criteria of completeness and correctness in making the transfer.

2. *Reasoning-gap activity*, which involves deriving some new information from given information through processes of inference, deduction, practical reasoning, or perception of relationships or patterns. One example is working out a teacher's timetable based on given class timetables. Another is deciding what course of action is best (for example cheapest or quickest) for a given purpose and within given constraints. The activity necessarily involves comprehending and conveying information, as in information-gap activity, but the information to be conveyed is not identical to that initially comprehended. There is a piece of reasoning which connects the two.

3. *Opinion-gap activity*, which involves identifying and articulating a personal preference, feeling, or attitude in response to a given situation. One example is story completion; another is taking part in the discussion of a social issue. The activity may involve using factual information and formulating arguments to justify one's opinion, but there is no objective procedure for demonstrating outcomes as right or wrong, and no reason to expect the same outcome from different individuals or on different occasions.

- **Pattison's Typology of Activities and Tasks**

This typology, which appeared nearly at the same time as the previous one, defined seven activity types (Nunan, 2004, pp.57-58):

1. *Questions and answers*: These activities are based on the notion of creating an information gap by letting learners make a personal and secret choice from a list of language items which all fit into a given frame (e.g. the location of a person or object). The aim is for learners to

discover their classmates' secret choices. This activity can be used to practice almost any structure, function or notion.

2. *Dialogues and role-plays*: These can be wholly scripted or wholly improvised. However, If learners are given some choice of what to say, and if there is a clear aim to be achieved by what they say in their role plays, they may participate more willingly and learn more thoroughly than when they are told to simply repeat a given dialogue in pairs.
3. *Matching activities*: Here, the task for the learner is to recognize matching items, or to complete pairs or sets. 'Bingo', 'Happy families' and 'Split dialogues' (where learners match given phrases) are examples of matching activities.
4. *Communication strategies*: These are activities designed to encourage learners to practice communication strategies such as paraphrasing, borrowing or inventing words, using gestures, asking for feedback and simplifying.
5. *Pictures and picture stories*: Many communication activities can be stimulated through the use of pictures (e.g. spot the difference, memory test, and sequencing pictures to tell a story).
6. *Puzzles and problems*: Once again, there are many different types of puzzles and problems. These require learners to make guesses, draw on their general knowledge and personal experience, use their imagination and test their powers of logical reasoning.
7. *Discussions and decisions*: These require the learner to collect and share information to reach a decision (e.g. to decide which items from a list are essential to have on a desert island).

- **Richards' Typology of Pedagogical Tasks**

A more recent typology was proposed by Richards (2001, p.162). He explained five (5) main types of pedagogical tasks:

1. *Jigsaw tasks*: These tasks involve learners in combining different pieces of information to form a whole (e.g. three individuals or groups may have three parts of a story and have to piece the story together).
2. *Information-gap tasks*: Tasks in which one student or group of students has one set of information and another student or group has a complementary set of information. They must negotiate and find out what the other party's information is in order to complete an activity.
3. *Problem-solving tasks*: Students are given a problem and a set of information. They must solve the problem. There is generally a single resolution of the outcome.
4. *Decision-making tasks*: Students are given a problem for which there are several possible outcomes and they must choose one through negotiation and discussion.
5. *Opinion exchange tasks*: Learners engage in discussion and exchange of ideas. They do not need to reach an agreement.

- **Ellis (1998) Classifications of Task Types**

Ellis (2011, pp. 213-214), categorized tasks based on their nature, aim and outcome into three main dichotomies:

- 1- Unfocused tasks versus focused tasks;
- 2- Input-providing versus output-prompting tasks;
- 3- Closed versus open outcome tasks.

Concerning the first category, Ellis explains that unfocused tasks engage learners in using the language communicatively, whereas focused tasks make them use the language to practice a particular linguistic structure but implicitly, which makes them different, according to him, different from 'situational grammar exercises'. The second classification of tasks is based on the different skills the learners are supposed to practice and develop being productive or receptive i.e. input providing tasks that engage learners in reading or listening while out-

prompting ones engage them in speaking or writing. He also notes that some tasks may engage the learners in two or more skills and he called these tasks ‘integrative ones’. The third and last categorization deals with the supposed outcome of the task being closed i.e. it has a limited number of outcomes or open, which means that there is an unlimited number of possible outcomes.

2.1.3.2.2.7.The Multi-syllabus Syllabus

This is sometimes referred to as the *mixed syllabus* or *the layered syllabus*, which can be either the *integration* or *combination* of the different types of syllabi discussed before (Krahnke, 1987). McDonough et al. (2013, p.13) point out that “It is, of course, unusual to find just one of these as the only organizing principle, in isolation from others”. This can be noticed when reading the table of contents of many textbooks as the reader would find the existence of not only a unique syllabus type. Brown (1995) explains this diversion by saying that it occurs in two ways: “sometimes two or more types of syllabuses may be mixed into what appears to be a different type of syllabus, and other times there may be secondary or tertiary syllabuses operating in layers underneath the primary syllabus” (p.12). The first way refers to a *mixed syllabus*, which can take different labels based on the syllabus types that are mixed such as *topical/task-based syllabus* or *situational-structural syllabus*. The second way refers to a *layered syllabus*, in which underneath the primary syllabus (it can be of either type), there is a secondary or maybe also a tertiary syllabus of any other type.

2.1.3.3.Materials (Textbooks) Analysis

Prior to deeply discussing the point of materials/textbooks analysis, it seems very necessary to make the distinction between the terms *analysis* and *evaluation*. In fact, analysis is a pre-evaluation stage. McGrath (2002) argues that:

At its most basic level, analysis is a process which leads to an objective, verifiable *description*. Evaluation, as the word suggests, involves *the making of judgements*...In its simplest form analysis seeks to discover what is there (Littlejohn, 1998), whereas evaluation is more concerned to discover whether *what one is looking for* is there – and, if it is, to put a value on it. (p.22)

In the next sections, the three levels of textbook analysis in addition to the three methods of textbook analysis and evaluation are presented.

2.1.3.3.1. Levels of Materials/Textbook Analysis

Littlejohn (1998) suggested three levels of analysis for materials and explained that these move from “making subjective selections of objective facts about the materials (level 1), through deductions about the demands likely to be made of teachers and learners (level 2), to conclusions about the apparent underlying principles and ‘philosophy’ of the materials (level 3)” (p.185). He summarised these levels in the following figure:

Figure 2.2.

Levels of analysis of language teaching materials (Littlejohn, 2011, p. 185)

1. ‘WHAT IS THERE’	<i>‘objective description’</i>
• Statements of description • Physical aspects of the material • Main steps in the instruction sections	
2. ‘WHAT IS REQUIRED OF USERS’	<i>‘subjective analysis’</i>
• Subdivision into constituent tasks • An analysis of tasks: what is the learner expected to do? • Who with? With what content?	
3. ‘WHAT IS IMPLIED’	<i>‘subjective inference’</i>
• Deducing aims, principles of selection and sequence • Deducing teacher and learner roles • Deducing demands on learner’s process competence	

Based on the classification done by Littlejohn, McGrath (2002) gave some examples of the features to be considered at each level. These are illustrated in Table 2.2.:

Table 2.2.

Textbook analysis at three levels based on Littlejohn (1998)

Note. Adopted from McGrath (2002, p.23)

Level	Focus of analysis	Examples of features to be considered
1	<i>‘what is there’</i>	Publication date; intended users; type of material; classroom time required; intended context of use; physical aspects, such as durability, components, use of colour, the way the material is divided up across components, how the student’s book is organised, and how the learners and teachers are helped to find their way around.
2	<i>‘what is required of users’</i>	Tasks: what the learner has to do; whether their focus will be on form, meaning or both; what cognitive operations will be required; what form of classroom organisations will be involved (e.g. Individual work, whole class); what medium will be involved; who will be the source of language or information.
3	<i>‘what is implied’</i>	Selection and sequencing of content (syllabus) and tasks; distribution of information across teacher and students components; reconsideration of information collected at level 1 and 2.

2.1.3.3.2. Analysis and Evaluation Methods

McGrath (2002) refers to three basic methods that are discerned by reviewing literature related to materials evaluation. These are *the impressionistic method*, *the checklist method*, and *the in-depth method*.

2.1.3.3.2.1. The Impressionistic Method

This method, as its name implies, is aimed at obtaining a general impression of the material (textbook). According to Cunningsworth (1995, p.1), “This kind of impressionistic overview is useful and gives us a general introduction to the material”. On the other hand, he argues that this overview cannot be relied on “to give enough detail to ensure a good match between what the coursebook contains and the requirements of the learning/teaching situation” (ibid). As an alternative, Cunningsworth referred to what he calls *in-depth evaluation*.

Yet, McGrath argues that the distinction between *impressionistic* and *in-depth methods* is not as clear as it might seem to be. To support this argument, he gave some examples in which an impression of the material can be gained by:

- looking at more *representative features* such as the design of a unit or a lesson or the treatment of particular language elements (Cunningsworth, 1995);
- through the analysis of exercises (Hutchinson, 1987);
- or through a combined approach by starting with a ‘guided browsing’ followed by an analysis of a single unit and an examination of how the language skills are treated across the book (Johnson, 1986).

2.1.3.3.2.2. The In-depth Method

In such a method, the materials analysis/evaluation tends to be more penetrating and has a specific agenda (Cunningsworth, 1995). If the impressionistic method seems to equate with Littlejohn’s (1998) level 1 discussed before, the in-depth method is likely to match the claims

for the materials in levels 2 and 3 (McGrath, 2002). In an in-depth evaluation, one “...could look at the balance of activities and skills contained in a unit, the potential for learner participation, the amount of new language introduced, the amount of recycling, etc.” (Cunningsworth, 1995, p. 2). According to him, contrary to the impressionistic method which he thinks is receptive, the in-depth method is characterised by its active nature as information is being collected based on a previously set agenda. In addition to that, Cunningsworth believes that a combination of the two methods will form a sound basis for the materials evaluation.

2.1.3.3.2.3. The Checklist Method

This method, as McGrath (2002) explains, is dealt with separately from the other methods, especially from *the impressionistic method*, because as the latter is characterised by its subjectivity, the first is *objective*. McGrath (2002, p.26) lists some advantages of the checklist method by saying that: “Compared to the most obvious alternatives, impressionistic evaluation involving dipping into a book and in-depth evaluation based on close analysis of features or sections, the checklist has at least four advantages”. Hence, he believes that this method is *systematic, cost-effective, explicit, and has a convenient data record format*. However, checklists, he adds, need to be adapted based on the purpose and the context in which they are being used.

2.2. Assessment/Evaluation

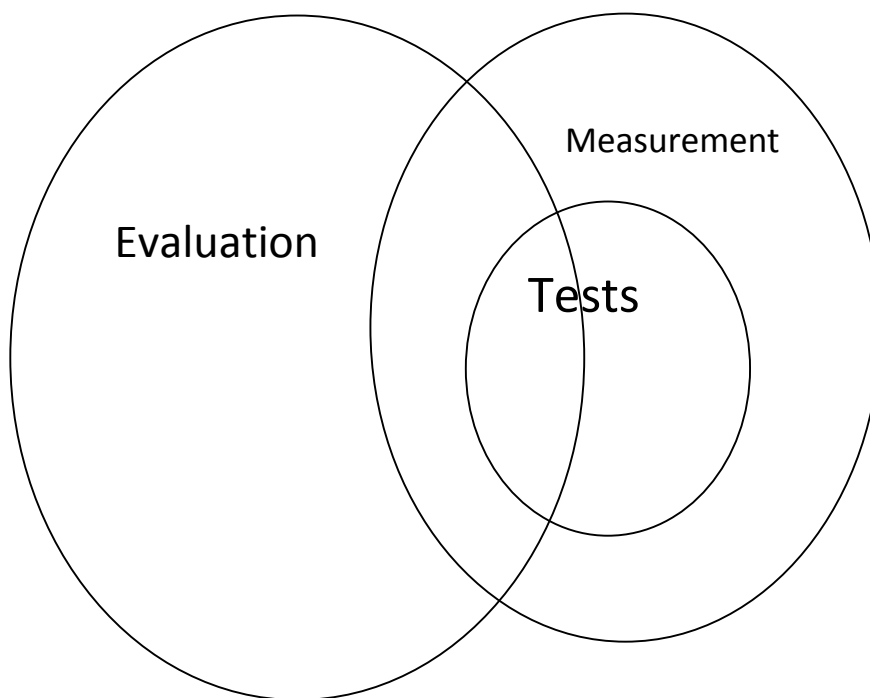
The term *Assessment* can be used interchangeably with *measurement*. It is “a procedure used to determine the degree to which an individual child possesses a certain attribute.” (ibid. p.6). Similarly, Bachman (1990) seems to use the term *measurement* instead of *assessment* while trying to clarify the distinctive characteristics of the terms “measurement”, “test” and “evaluation”; terms which, he admits that “are often used synonymously; indeed they may, in practice, refer to the same activity.” (p.18).

The relationship between the three terms, according to Bachman, can be explained as follows “not all measures are tests, not all tests are evaluative and not all evaluation involves either measurement or tests.” (ibid, p.24). The figure below illustrates this relationship:

Figure 2.3.

Relationships among measurement, tests and evaluation

Note. Adapted from Bachman (1990, p.23)



2.2.1. Types of Assessment

Mihai (2010) provides a comprehensive categorisation of assessment based on its instruments and procedures. Assessment, then, is categorised according to its: *intention, purpose, interpretation, and administration*. Table 2.3., summarises the types of assessment that belong to each category:

Table 2.3.

The Categories and Types of Assessment

Note. Adopted from Mihai (2010, p.27)

The Categories and Types of Assessment	
Category of Assessment	Type of Assessment
Intention	Informal
	Formal
Purpose	Formative
	Summative
Interpretation	Norm-referenced
	Criterion-referenced
Administration	Classroom-based
	Large scale

2.2.1.1. Informal/Formal Assessment

Teachers often tend to check learners' understanding, comment on a mispronunciation of a word or a piece of writing (paragraph or essay), praise learners' answers, and so on without prior planning. These are examples of *informal assessment*, which can, according to Brown (2004):

...take a number of forms, starting with incidental, unplanned comments and responses, along with coaching and other impromptu feedback to the student... Informal assessment does not stop there. A good deal of a teacher's informal assessment is embedded in classroom tasks designed to elicit performance without recording results and making fixed judgments about a student's competence. (p.5)

On the other hand, *formal assessment* is carried out with the prior intention to measure the extent to which a particular learning objective (s) has been achieved. This type of assessment is planned beforehand, it is systematic and learners know that they are being assessed well in advance. Generally, formal assessment is scored; be it a test, a quiz, an examination or even in the form of a portfolio, student blog or oral presentation. These latter forms of assessment make a clear cut to understanding the relationship between *formal assessment* and *testing*. Brown (2004) explains it in the way that "all tests are formal assessments but not all formal assessment is testing" (p.6).

2.2.1.2. Formative/Summative Assessment

Based on its purpose, assessment in education is classified into *formative* and *summative*, which are, often, referred to as *assessment for learning* and *assessment of learning* respectively (Black & William, 1998; Jones, 2005; Everhard & Murphy, 2015).

Black & William (1998) confirm that the focus of assessment studies shifted “towards greater interest in the interactions between assessment and classroom learning and away from concentration on the properties of restricted forms of test which are only weakly linked to the learning experiences of students.” (p.7). The authors mean that the focus of attention shifted from concentrating on summative assessment, which is concerned with the final product or achievement towards formative assessment (or assessment *for* learning). This latter they add:

...does not have a tightly defined and widely accepted meaning. In this review, it is interpreted as encompassing all those activities undertaken by teachers, and/or by their students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged. (ibid, pp. 7-8)

2.2.1.3. Norm-referenced/Criterion-referenced Assessment

In norm-referenced assessment, the students’ scores are reported back to test-takers in the form of a numerical score (for example 210 out of 250) and are compared to the scores of a reference group which is considered to be the norm. The results of the norm group, which is supposed to have the same characteristics as the students being assessed, are then used as reference points to interpret the performance of those students. Mihai (2010) listed these norms as follows:

the mean (the average score), the median (the middle score), the mode (the most common score), the standard deviation (how varied and spread out the group scores are), and the percentile rank (an indication of how well a particular test taker has performed compared to other group members; for example, a 78 percentile means that the test taker’s score was higher than 78% of the students taking the test, but lower than 22% of the students taking the test, that day). (pp. 29-30)

Mihai further explains that these assessments do not reveal any information about a student's mastery of the course content. However, it only ranks the students' scores according to their peers' performance.

On the other hand, according to Brown (2004, p.7), criterion-referenced tests “are designed to give test-takers feedback, usually in the form of grades, on specific course or lesson objectives”. Hence, contrary to norm-referenced tests in which students are evaluated in comparison to one another's performance, criterion-referenced tests' scores will indicate whether the students have attained a certain degree of academic content-based proficiency (Mihai, 2010).

Table 2.4., adopted from Brown (1996, p.3), illustrates the main differences between Norm-referenced and criterion-referenced tests:

Table 2.4.

Differences between Norm-referenced and Criterion-referenced Tests

Characteristic	Norm-referenced	Criterion-referenced
Type of Interpretation	Relative (A student's performance is compared to that of all other students in percentile terms.)	Absolute (A student's performance is compared only to the amount, or percentage, of material learned.)
Type of Measurement	To measure general language abilities or proficiencies	To measure specific objectives-based language points
Purpose of Testing	Spread students out along a continuum of general abilities or proficiencies	Assess the amount of material known, or learned, by each student

Distribution of Scores	Normal distribution of scores around a mean	Varies, usually non-normal (students who know all of the material should all score 100%)
Test Structure	A few relatively long subtests with a variety of question contents	A series of short, well-defined subtests with similar question contents
Knowledge of Questions	Students have little or no idea what content to expect in questions	Students know exactly what content to expect in test questions

2.2.1.4. Classroom-based/Large-scale Assessment

This classification is done based on the administration of the assessment tool. Therefore, a classroom-based assessment is designed by the teacher and administered in the classroom as a formative assessment that takes place regularly and it can take several forms such as a test at the end of a learning unit, portfolios, project work, student's self/peer evaluation ...etc.

Large-scale assessment, on the other hand, generally takes the form of a standardized test and is designed to be administered to a large population of students/test takers. This kind of assessment includes national/state tests and tests administered at the end of the course.

2.2.2. Test Types

2.2.2.1. Placement Test

Hughes (1989, p.14) defines a placement test as one that is "intended to provide information that will help to place students at the stage (or in the part) of the teaching programme most appropriate to their abilities. Typically, they are used to assign students to classes at different levels". Very often, placement tests are bought by institutions from assessment centres which design them for commercial purposes. However, this is not something

right to do as there is no standard test that can always categorize students applying for different courses and following varied curricula. In order to achieve the ultimate objective of a placement test, which is to correctly place a student into a course or a level, the performance of a student on the test will indicate the point at which he/she finds the material appropriately challenging i.e. neither too easy nor too difficult (Brown, 2004).

2.2.2.2.Proficiency Test

This kind of test is designed to measure someone's overall ability (Hughes, 1989) or competence (Brown, 2004) in a language without taking into account whether they had previous training in that language or not. According to Brown (2004, pp.44-45), "Proficiency tests are almost always summative and norm-referenced...And because they measure performance against a norm, with equated scores and percentile ranks taking on paramount importance, they are usually not equipped to provide diagnostic feedback".

For some proficiency tests, to be 'proficient' means to have a command/mastery of a language to be used for *a particular purpose* (Hughes, 1989). Examples of standardized proficiency tests include the Test of English as a Foreign Language (TOEFL), and the International English Language Testing System (IELTS). These tests are recognized by thousands of universities and institutions around the world as evidence of a student's prospective ability to register for a course or undertake research work in an English-speaking environment.

2.2.2.3.Diagnostic Test

As their name suggests, diagnostic tests are used to identify a learner's strengths and weaknesses in particular aspects of the language. Some placement tests can be used as diagnostic tests if they are able to provide information well beyond the course level a student is

to be designated (Brown, 2004). Brown (1996) has clarified a slight difference between diagnostic tests and achievement tests, which is worth citing:

While achievement decisions are usually focused on the degree to which the objectives have been accomplished at the end of the program or course, diagnostic decisions are normally made along the way as the students are learning the language. As a result, diagnostic tests are typically administered at the beginning or in the middle of a language course. (p.15)

Hence, a diagnostic test offers more detailed information about an individual learner's deficiencies for the sake of bridging the gap by including them in the course the students will take or correcting them in case the test is conducted in the middle of the learning process.

2.2.2.4.Achievement Test

Harmer (2007, p.380) considers that “these tests are designed to measure learners’ language and skill progress in relation to the syllabus they have been following”. Two main types of achievement tests are found in literature: *final (summative) and progress (formative) achievement tests*. As its name reveals, the first type is generally written and administered by official examination boards, ministries of education or members of teaching institutions (Hughes, 1989) within a particular time frame at the end of a course of study. The second one “will offer washback about the quality of a learner’s performance in subsets of the unit or course. This washback contributes to the formative nature of such tests” (Brown, 2004, p. 48).

Accordingly, Brown (2004) listed some factors that determine the specifications of an achievement test which are:

- The objectives of the lesson, unit, or course being assessed,
- The relative importance (or weight) assigned to each objective,
- The tasks employed in classroom lessons during the unit of time,
- Practicality issues, such as the time frame for the test and turnaround time, and

- The extent to which the test structure lends itself to formative washback.

2.2.3. Framework of Language Test Task's Characteristics

Bachman (1996) provided a framework (Figure 2.3.) specifying the characteristics of a language task consisting of a set of features that describe the following aspects of tasks: setting, test rubric, input, expected response, and the relationship between input and response. The purpose of this framework, as Bachman (1996, p.47) puts it, is “to provide a basis for language development and use”. He adds that this involves three activities:

- 1- Describing TLU (Target Language Use) tasks as a basis for designing language test tasks,
- 2- Describing different test tasks in order to ensure their compatibility, and as a means for assessing reliability, and
- 3- Comparing the characteristics of TLU and test tasks to assess authenticity.

Bachman further clarifies that this framework can be applied to both test tasks and TLU tasks that are relevant to designing and developing test tasks. Hence, this framework can be used to analyse both testing and instructional tasks, which makes it quite similar in many aspects to the schedule for task analysis by Littlejohn (2011) (See Appendix A).

Figure 2.4.

Language-test Task Characteristics (Bachman, 1996, pp.49-50)

Task Characteristics
Characteristics of the setting <i>Physical characteristics</i> <i>Participants</i> <i>Time of task</i>
Characteristics of the test rubrics Instructions Language (native, target) Channel (aural, visual) Specification of procedures and tasks Structure Number of parts/tasks Salience of parts/tasks Sequence of parts/tasks Relative importance of parts/tasks Number of tasks/items per part Time allotment Scoring method Criteria for correctness Procedures for scoring the response Explicitness of criteria and procedures
Characteristics of the input Format Channel (aural, visual) Form (language, non-language, both) Language (native, target, both) Length Type (item, prompt) Degree of speededness Vehicle (live, reproduced, both) Language of input Language characteristics Organisational characteristics Grammatical (vocabulary, syntax, phonology, graphology) Textual (cohesion, rhetorical/conversational organizational) Pragmatic characteristics Functional (ideational, manipulative, heuristic, imaginative) Sociolinguistic (dialect/variety, register, naturalness, cultural reference, and figurative language) Topical characteristics

Characteristics of the expected response**Format**

- Channel (aural, visual)
- Form (language, non-language, both)
- Language (native, target, both)
- Length
- Type (selected, limited production, extended production)
- Degree of speededness
- Vehicle (live, reproduced, both)

Language of expected response

Language characteristics

Organisational characteristics

- Grammatical (vocabulary, syntax, phonology, graphology)

- Textual (cohesion, rhetorical/conversational organizational)

Pragmatic characteristics

- Functional (ideational, manipulative, heuristic, imaginative)

- Sociolinguistic (dialect/variety, register, naturalness, cultural reference, and figurative language)

Topical characteristics

Relationship between input and response

- Reactivity (reciprocal, non-reciprocal, adaptive)

- Scope of relationship (broad, narrow)

- Directness of relationship (direct, indirect)

2.2.4. Matching Tests to Decision Purposes

Table 2.5., summarises the test qualities and the purpose lying behind before taking a decision to use any of the four test types discussed in the previous point.

Table 2.5.*Matching Tests to Decision Purposes (Brown, 1996, p.9)*

Test Qualities	Type of Decision			
	Norm-referenced		Criterion-referenced	
	Proficiency	Placement	Achievement	Diagnostic
Detail of Information	Very General	General	Specific	Very specific
Focus	Usually, general skills prerequisites to entry	Learning points all levels and skills of program	Terminal objectives of course or program	Terminal and enabling objectives of course
Purpose of Decision	To compare individual overall with other groups/individuals	To find each student's appropriate level	To determine the degree of learning for advancement or graduation	To inform students and teachers of objectives needing more work
Relationship to program	Comparisons with other institutions	Comparisons within program	Directly related to objectives of program	Directly related to objectives still needing work
When Administered	Before entry and sometimes at exit	Beginning of program	End of courses	Beginning and/or middle of courses
Interpretation of Score	Spread of scores	Spread of scores	Number and amount of objectives learned	Number and amount of objectives learned

2.2.5. Cardinal Criteria for ‘Testing a Test’

Validity, reliability, practicality, equivalency and wash-back are of paramount importance while designing language tests in particular and are, also, major principles to be considered in language assessment as a whole.

2.2.5.1. Validity

The ‘validity’ of a test is perhaps the most important and at the same time, the most complex feature to be achieved. Although research about test validity, in general, makes “...investigations into the extent to which a test measures what it is supposed to measure” (Palmer, 1981, p.135), measurement specialists have, traditionally, discussed different types of validity. Validity which is seen by some researchers as a ‘unitary concept’ (Messick, 1989; Bachman, 1990) encompasses four main distinct types. These are discussed in the following elements.

2.2.5.1.1. Content Validity

One of the most important types of test validity is, generally, referred to as ‘content validity’ (Alderson & Hughes, 1981), which has two aspects: ‘content relevance’ and ‘content coverage’ (Bachman, 1990) and in some references, it is given the term ‘content-related evidence’ (Brown, 2004). Kubiszyn & Borich (2013) consider that:

The content validity evidence for a test is established by inspecting test questions to see whether they correspond to what the user decides should be covered by the test... In the context of classroom testing, content validity evidence answers the question “Does the test measure the instructional objectives?” In other words, a test with good content validity evidence matches or fits the instructional objectives. In a high-stakes testing context, content validity evidence answers the question “Does the test align with state academic standards (or the Common Core State Standards)?” (p.327)

They also raise two major problems with content validity. The first is when the concept being tested is a personality trait for example, where it is difficult to anticipate beforehand what a suitable question looks like. The second problem is when the test *looks* valid but does not really measure what it is meant to measure as the reading level is high or the test items are poorly constructed (ibid).

2.2.5.1.2. Criterion-related Validity

Criterion-related validity or criterion-related evidence is defined by Brown (2004, p.24) as “the extent to which the ‘criterion’ of the test has actually been reached”. Accordingly, Brown classifies criterion-related validity into two categories: The first is *concurrent validity*, which refers to the test results’ reflection of the same level of performance in other concurrent activities or learning tasks. The second category is *predictive validity* in which the criterion of the assessment is to predict the test taker’s likelihood to succeed in what the test opens the door for; that is predictive validity is most needed in placement tests, language aptitude tests, and entry tests for example.

2.2.5.1.3. Construct Validity

Prior to explaining what construct validity is meant to deal with, it is necessary to understand the sense of ‘*construct*’ as a term. A *construct*, according to Brown (2004), is any theory, hypothesis, or model that makes attempts to explain phenomena observed in our universe of perceptions. Examples of constructs are *oral proficiency* and *communicative language use*.

Bachman (1990, pp. 254-255) considers that “Construct validity concerns the extent to which performance on tests is consistent with predictions that we make on the basis of a theory of abilities, or constructs.” Messick (1994) further explains what construct validity refers to by saying that:

In principle as well as in practice, construct validity is based on an integration of any evidence that bears on the interpretation or meaning of the test scores including content- and criterion-related evidence, which are thus subsumed as part of construct validity. (p.3)

Kubiszyn & Borich (2013) further clarify that *construct validity* is different from the two categories of criterion-related validity evidence as they consider that “It differs from concurrent validity evidence in that there is no accepted second measure available of what you’re trying to measure, and it differs from predictive validity evidence in that there is no available measure of future behavior” (p.329).

In the domain of assessment, *construct validity*, then, asks the question: Does this test actually tap into the theoretical construct as it has been defined? (Brown, 2004). Hence, according to Mihai (2010, p.39), “Test designers and classroom teachers should consult state and national standards when creating large-scale or classroom-based assessments”.

2.2.5.1.4. Face Validity

In addition to the three major types of validity discussed before, another type that is not as important is *face validity*, which as Mousavi (2002) puts it:

...refers to the degree to which a test *looks* right, and *appears* to measure the knowledge or abilities it claims to measure, based on the subjective judgement of the examinees who take it, the administrative personnel who decide on its use, and other psychometrically unsophisticated observers. (p.244)

Face validity is, therefore, based on the test-takers or maybe the test-givers’ perceptions about whether the test, on the face of it, tests what it is designed to test or not. A language test has *face validity* if it appears, at a glance, to be measuring students’ knowledge of aspects of the language being tested and not measuring knowledge of history or geology concepts. It is clear

that this type of validity cannot be measured by using empirical instruments as it is dependent on intuitive perceptions.

2.2.5.2. Reliability

Reliability is often considered as consistency in measurement (Bachman & Palmer, 1996). According to them, “A reliable test score will be consistent across different characteristics of the testing situation” (p.19). This means that a test is said to be *reliable* if an individual takes the same test under different conditions and in different settings but gets similar scores. Another case is that when two versions of a test are designed to be used interchangeably, their outcomes should be similar for the test to be considered *reliable*. To support this, Mihai (2010, p.42) points out that “a reliable test is a consistent measure that stays consistent regardless of the time, the test format, raters, and other characteristics of the testing environment”. Brown (2004) discussed four different types of reliability: *student-related reliability, rater reliability, test administration reliability, and test reliability*.

2.2.5.2.1. Student-related Reliability

Test-takers can affect the reliability of a test through some factors related to them such as illness, fatigue, anxiety, family/personal problems among other psychological factors. In order to establish student reliability, it is necessary to ensure that the test-takers are in a good health and stable psychological state and that there are no such reasons that can distract their performance on the test.

2.2.5.2.2. Rater Reliability

This is closely related to the scoring process and the item types in a test. This type of reliability is not an issue in MCQs, true or false, and other test items, in which only one answer is accepted as a correct one. On the contrary, this becomes serious in short-answer or essay items where the rater does not have a list of exact correct answers and the test scores' reliability

can be affected by either the raters' subjectivity, human errors due to inattention, fatigue or lack of experience for example, or even the use of different scoring standards by the two or three scorers. Such a bias can be classified into two categories: *interrater bias* (in the case of two or more scorers) and *intrarater bias* (when one person scores the tests). Rater reliability is harder to achieve especially in writing and speaking tests as they consist of a complex set of traits that are difficult to define and to defragment. This necessitates a scoring instrument similar to a checklist to be used by the rater(s) in a trial to stick to the scoring standards.

2.2.5.2.3. Test Administration Reliability

This has to do with the overall conditions underlying test administration and taking. A test is said to achieve *test administration reliability* if the assessment conditions are the same for all test takers and the test administration is consistent. The conditions referred to here may include changes in the temperature, the noise inside/outside the test room, unclear test sheets, poor lighting...etc. The same chances of having the best test administration conditions should be given to test-takers sitting for the same test in different periods to achieve this type of reliability.

2.2.5.2.4. Test Reliability

The test in itself can cause measurement errors due to its format or content. In terms of format, the test can be too long and this leads either to test-takers losing concentration as they reach the last items/exercises of the test (due to fatigue) or the lack of time to answer these items/exercises by slow learners in the case of timed tests. Sometimes, the items themselves are poorly written, which makes them act as a distractor that results in raising the test-takers levels of anxiety and negatively affects their performances on the test.

2.2.5.3. Practicality

In addition to being valid and reliable, a test has to be practical. Bachman (2004) limits the meaning of a practical test in:

- being not excessively expensive,
- staying within appropriate time constraints,
- it is relatively easy to administer, and
- having a scoring/evaluation procedure that is specific and time efficient.

Hence, it is necessary to make a certain balance between validity and practicality while designing tests. Mihai (2010) illustrates this by saying:

The best way to assess a certain skill is to directly test that skill. If we want to assess academic writing, the most valid method of assessment is to actually have students put together a portfolio that contains a selection of the students' most representative writing samples in the content area. However, such an assessment instrument is problematic because of cost and time... Such an assessment may be used at the classroom level, but for a state or a national assessment, a different instrument would have to be employed. (p.44)

2.2.5.4.Equivalency

If a test is directly based on curricular standards or instructional activities, this means that it has the property of *equivalency* (Mihai, 2010). In the classroom context, this refers to the influence that teaching has on a test or other assessment's design. In state-wide examinations, *equivalency* is achieved when an exam is based on some standards that are related to the content area and these are set by the testing organisation or institution. Mihai (2010, p.46) further explains that "the teachers are basing their instruction on state standards, and the educational materials are based on these standards as well."

2.2.5.5.Washback

This is sometimes referred to as *backwash* (Hughes, 1989). Hughes defines it as "The effect of testing on teaching and learning" (ibid, p.1). According to him, washback (or backwash) can be beneficial or harmful. It is beneficial if a test is considered to be important and then the preparation for it will be the focus of all teaching and learning. In case the test

content is different from the prior instruction and its objectives, then the washback tends to be harmful.

Brown (2004, p.28) considers that “In large-scale assessment, washback generally refers to the effects the tests have on instruction in terms of how students prepare for the test. ‘Cram’ courses and ‘teaching to the test’ are examples of such washback”. These examples given by Brown illustrate the negative side of washback. This latter can have also a positive side when, for example, teachers give praise or extensive comments on test performances and do not satisfy themselves by giving grades or scores solely.

Conclusion

The discussion in this chapter has been centred on some basic curriculum components, namely syllabus design, materials development and assessment and evaluation. It was necessary to define the key terms each time to limit the reader’s understanding and conception each time a reference to one of them. Throughout the discussion of the last component, i.e. assessment and evaluation, a clear distinction has been made between the two terms. In addition, *testing*, being a form of assessment, has been dealt with in a relatively detailed manner due to its importance vis-à-vis the nature of the present research. For the same reason, more focus and concern are to be put on a specific type of test, which is the *achievement test*, which is the topic of chapter three.

Chapter Three:

A Focus on Achievement Tests

Table of Contents

Introduction	76
3.1. Nature and Features of Achievement Tests	76
3.2. Functions of Achievement Tests.....	77
3.3. Limitations of Achievement Tests	77
3.4. Taxonomies of Educational Objectives and Achievements.....	78
3.4.1. Bloom’s Taxonomy	78
3.4.2. Anderson and Krathwohl's Taxonomy	81
3.4.3. Marzano and Kendall’s New Taxonomy	84
3.5. Achievement Test Planning and Design	85
3.5.1. Setting the Purpose for Testing	85
3.5.2. Developing a Test Plan.....	86
3.5.3. Selection of Item Formats	86
3.5.3.1. Essay Tests	86
3.5.3.2. Objective Tests	88
3.5.4. Quantity of Test Items	93
3.5.5. Distribution of Test Items’ Difficulty Level.....	94
3.6. Describing and Summarizing Achievement Test Results.....	95
3.6.1. Summarizing Data.....	95
3.6.1.1. Frequency Distribution.....	95
3.6.1.2. Graphing Data	99
3.6.2. Describing and Summarizing Score Distributions	102
3.6.2.1. Measures of Central Tendency.....	102
3.6.2.2. Measures of Variability	105
3.6.2.3. Normal Distribution.....	108
3.6.2.4. Standard Scores	110
Conclusion.....	111

Introduction

An achievement test evaluates accomplishment or the degree of learning that has taken place (Cohen & Swerdlik, 2009). It has some functions to fulfil as well as limitations of use. In this chapter, in addition to presenting an achievement test's functions and limitations, a set of the most pertinent taxonomies of educational objectives are discussed. Moreover, an emphasis is put on the planning and design of an achievement test, some statistical basics of the different ways of summarising and interpreting a test's scores and also the description of scores' distributions. A reference to standardised achievement tests is made at the end of the chapter.

3.1.Nature and Features of Achievement Tests

Bachman (1990) refers to achievement tests as *syllabus-based tests*. In their classification of tests in terms of the information they provide, Genesse and Upshur (1996, p.153), says that "Achievement tests provide information about student attainment in relation to instructional objectives or a defined domain of language". A similar specification of achievement tests was given by Brown (2004, p. 47) by saying:

An achievement test is related directly to classroom lessons, units, or even a total curriculum. Achievement tests are (or should be) limited to particular material addressed in a curriculum within a particular time frame and are offered after a course has focused on the objectives in question.

Moreover, Brown (2004, p.48) discussed some of the features of achievements, and these are listed below:

- Achievement tests are often summative as they are administered at the end of a unit or term of study.
- They can also play an important formative role by providing washback about the quality of learners' performances in parts of the unit of course.

- They range from five-ten minutes quizzes to three-hour final examinations.
- They can have an infinite number of formats with an infinite variety of item types.

3.2.Functions of Achievement Tests

Ebel and Frisbie (1991), in their discussion of the functions that an achievement test can serve, listed three main elements. These are summarised as follows:

- The contribution to the evaluation of educational progress and attainment by assigning meaningful and accurate grades.
- Motivating and directing students' learning.
- The process of taking a test or an examination and discussing its outcomes can serve as a rewarding experience for the students as any other life activity they engage in.

3.3.Limitations of Achievement Tests

As with any piece of work, performed task or achieved mission, an educational achievement test has some limitations that were discussed again by Ebel and Frisbie (1991). The main points they evolved in this section are the following:

- Unlike physical measurement, the best mental measurement can do is to provide an approximate rank order of learners in terms of their ability to accomplish a more or less well-defined set of activities.
- Some intangible outcomes cannot be measured by a test as they represent subtle and unforeseen effects on an individual learner by the educational process. These may appear long after the student leaves school and may be of crucial importance on the individual's level.
- Paper and pencil tests are used for measuring verbal and numerical knowledge and understanding as well as solving problems of that kind but they cannot test physical development or social skills attainment.

3.4. Taxonomies of Educational Objectives and Achievements

According to what was previously mentioned, Brown (2004) listed the learning objectives of the lesson, unit, or course being assessed at the top as a contributor to the definition of achievement test specifications. Some educators worked on clearing the ambiguity related to the statement of these instructional objectives and translating them into test items. As a result, a number of taxonomies were set up. Among the most widely discussed taxonomies in literature are Bloom (1956), Anderson and Krathwohl (2001) and Marzano and Kendall (2007).

3.4.1. Bloom's Taxonomy

Bloom and others (1956) categorized the learning outcomes into three different domains: *cognitive*, *affective* and *psychomotor*. Yet, the *cognitive domain* was given priority by dealing with it in Handbook I of their work as it was clearly stated:

Our original plans called for a complete taxonomy in three major parts – the cognitive, the affective and the psychomotor domains. The cognitive domain, which is the concern of this Handbook, includes those objectives which deal with the recall or recognition of knowledge and the development of intellectual abilities and skills. This is the domain which is most central to the work of much current test development. It is the domain in which most of the work in curriculum development has taken place and where the clearest definitions of objectives are to be found phrased as descriptions of student behaviour. (p.7)

The classification of educational objectives and test exercises according to this taxonomy is done based on learners' cognitive behaviours dividing them into six major classes and each class has sub-categories as follows:

1.00 Knowledge

- 1.10 Specifics

1.11 Terminology

1.12 Facts

- 1.20 Ways and means of dealing with specifics

1.21 Conventions

1.22 Trends and sequences

1.23 Classification and categories

1.24 Criteria

1.25 Methodology

- 1.30 Universals and abstractions

1.31 Principles and generalizations

1.32 Theories and structures

2.00 Comprehension

- 2.10 Translation
- 2.20 Interpretation
- 2.30 Extrapolation

3.00 Application

4.00 Analysis

- 4.10 Analysis of elements
- 4.20 Analysis of relationships
- 4.30 Analysis of organisational principles

5.00 Synthesis

- 5.10 Production of a unique communication
- 5.20 Production of a plan, or proposed set of operations
- 5.30 Derivation of a set of abstract relations

6.00 Evaluation

- 6.10 Judgements in terms of internal evidence

- 6.20 Judgements in terms of external criteria

The second part of the taxonomy, which was discussed in *Handbook II: The Affective Domain* by Krathwohl et al. (1964) cited in Bloom (1956), relates to "... objectives which describe changes in interest, attitudes, and values, and the development of appreciations and adequate adjustment" (p.7).

Although Bloom and his colleagues recognise the existence of the manipulative or motor-skill area, they admit that they found little done about it in secondary schools and colleges. Hence, they didn't find it so useful to develop a classification of objectives concerning this domain at that time. However, Ebel and Frisbie (1991, p.52) mentioned some taxonomies that dealt with the psychomotor domain. These were developed by educational psychologists such as Harrow (1972) and Kibler et al. (1981). It is clarified that "The categories of each system vary but generally relate to gross bodily movements, finely coordinated bodily movements, nonverbal communications, and speech behaviours" (ibid).

The taxonomy of Bloom, being a founding framework, received a lot of criticism. Marzano and Kendall (2007, p.8) clarified that according to Furst (1994), "One of the most common criticisms was that the taxonomy oversimplified the nature of thought and its relationship to learning". They even add that the problems with the taxonomy were implicitly acknowledged by its authors. These problems with the taxonomy's structure were basically discussed at the analysis and evaluation levels. The following extracts from Bloom et al. (1956) were given as evidence:

- "It is probably more defensible educationally to consider analysis as an aid to fuller comprehension (a lower class level) or as a prelude to an evaluation of the material" (p. 144).
- "Although evaluation is placed last in the cognitive domain because it is regarded as requiring to some extent all the other categories of behavior, it is not necessarily the last

step in thinking or problem-solving. It is quite possible that the evaluation process will in some cases be the prelude to the acquisition of new knowledge, a new attempt at comprehension or application, or a new analysis and synthesis” (p. 185).

Then, many subsequent revisions to this taxonomy were conducted and alternative frameworks were suggested in a trial to overcome the problems and the conceived shortcomings of this initial work. Two of these taxonomies are by Anderson and Krathwohl (2001), and Marzano and Kendall (2007).

3.4.2. Anderson and Krathwohl Taxonomy

This is a revised version of Bloom’s taxonomy of educational objectives. Appendix A of the book by Anderson and Krathwohl (2001), summarised the changes applied to the original taxonomy. It is mentioned that:

Readers familiar with the original framework will recognize that we have made a number of changes, 12 in all: four changes in emphasis, four in terminology, and four in structure. Most important, we have changed the focus of the document. (ibid, p.263)

These changes are summarised as follows:

- Changes in Emphasis

- 1- The revision’s primary focus is on the taxonomy in use:** Contrary to the original focus on assessment and also on providing examples of test items for the six categories, the focus shifts to the use of the revised taxonomy in curriculum planning, assessment and instruction and the alignment of all three together.
- 2- The revision is aimed at a broader audience, emphasizing teachers:** Teachers of elementary and secondary education are also kept in mind while designing the taxonomy and examples drawn from these levels predominate the revision.

3- Sample assessment tasks are included primarily to convey meaning:

4- The revision emphasizes the subcategories

- Changes in Terminology

5- Major category titles were made consistent with how objectives are framed

6- The knowledge subcategories were renamed and reorganized

7- Subcategories of the cognitive process categories were replaced by verbs

8- Comprehension and synthesis were retitled

- Changes in Structure

9- The noun and verb components of objectives became separate dimensions

10- The two dimensions are the basis for the analytical tool i.e. the taxonomy table

11- The process categories do not form a cumulative hierarchy

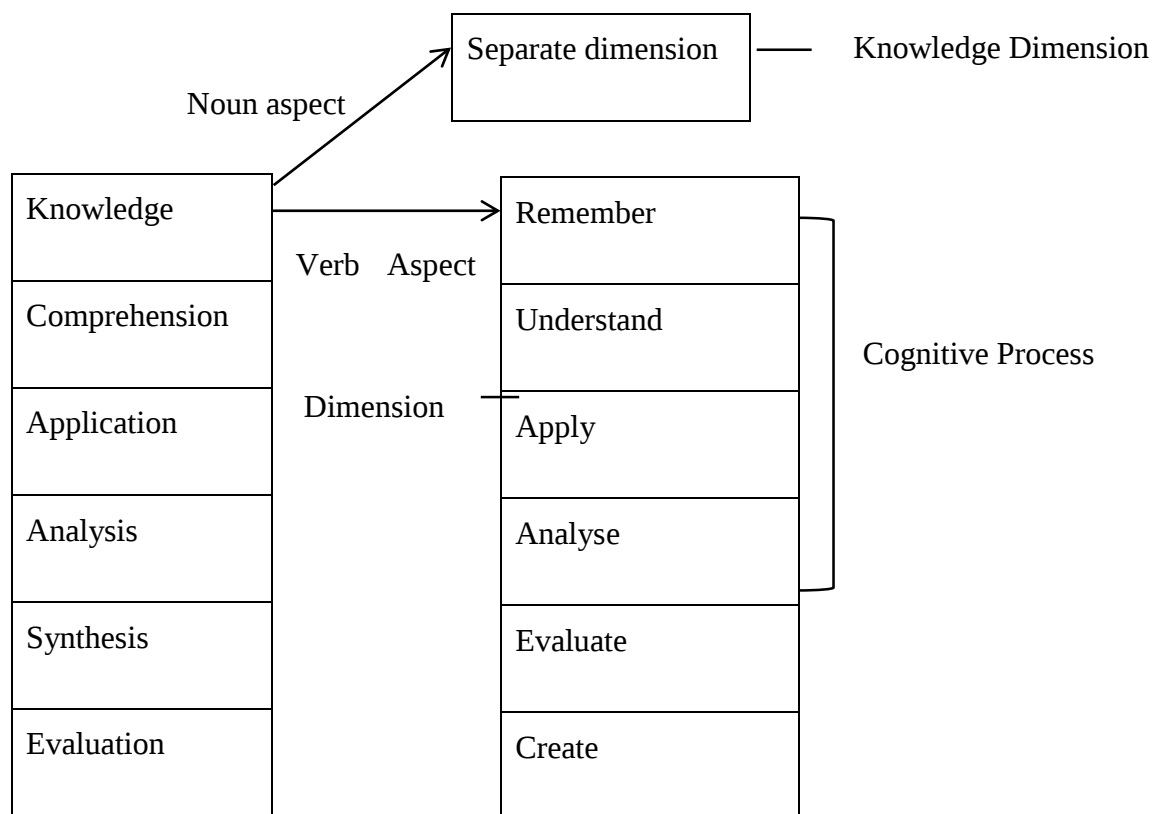
12- the order of synthesis/create and evaluation/evaluate was interchanged

These structural changes from the original framework (Bloom's Taxonomy) are summarised in figure 3.1.:

Figure 3.1.

Summary of the Structural Changes from the Original Framework to the revision

Note. Adopted from Anderson and Krathwohl (2001, p. 268)



The two-dimensional nature of this taxonomy was justified by Anderson et al. (2001) by saying that:

In our Taxonomy we are classifying objectives. A statement of an objective contains a verb and a noun. The verb generally describes the intended cognitive process. The noun generally describes the knowledge students are expected to acquire or construct. Consider the following example: "The student will learn to distinguish (the cognitive process) among confederal, federal, and unitary systems of government (the knowledge). (p.4)

Unlike the cognitive process dimension which kept its six categories division, the knowledge dimension is partitioned into four categories: *Factual*, *Conceptual*, *Procedural*, and *Metacognitive* knowledge. These are classified in a continuum from *concrete* to *abstract*.

3.4.3. Marzano and Kendall's New Taxonomy

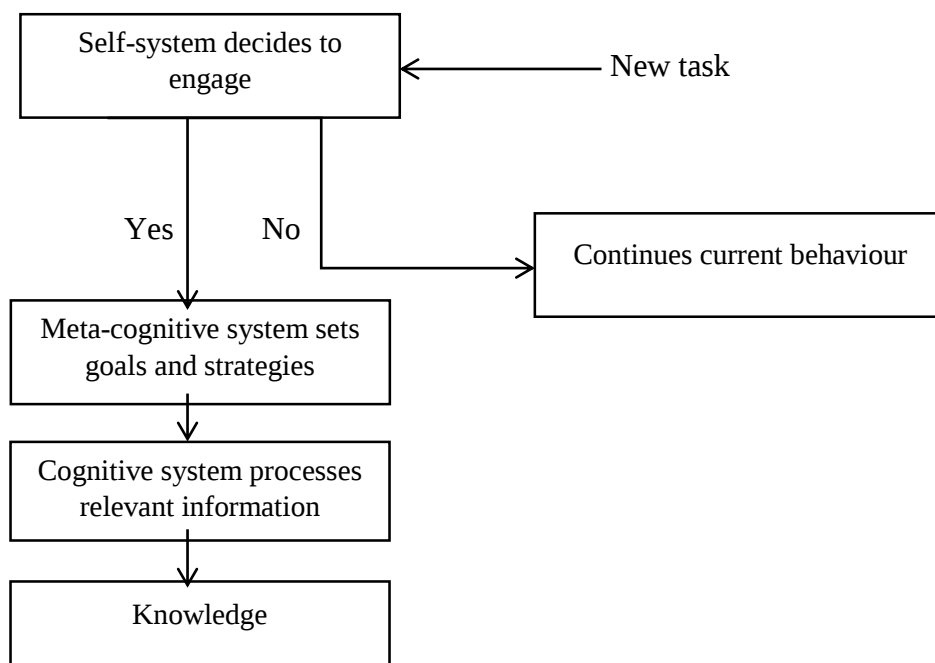
In their explanation of the theoretical basis of their taxonomy referred to as *The New Taxonomy of Educational Objectives*, Marzano and Kendall (2007) seem to build it by trying to avoid the weaknesses that Bloom's taxonomy, they assume, suffers from. One of these latter is the classification of the levels in Bloom's taxonomy based on the degree of difficulty. Moreover, they explained some of the differences between their taxonomy and the ones previously presented in this chapter i.e. Bloom (1956) and Anderson and Krathwohl (2001).

Marzano and Kendall suggested a model on which they have relied to make the design of their new taxonomy. This model is based on three mental systems in addition to knowledge as a separate dimension. This is illustrated in figure 3.2.:

Figure 3.2.

Model of Behaviour

Note. Adopted from Marzano and Kendall (2007, p. 11)

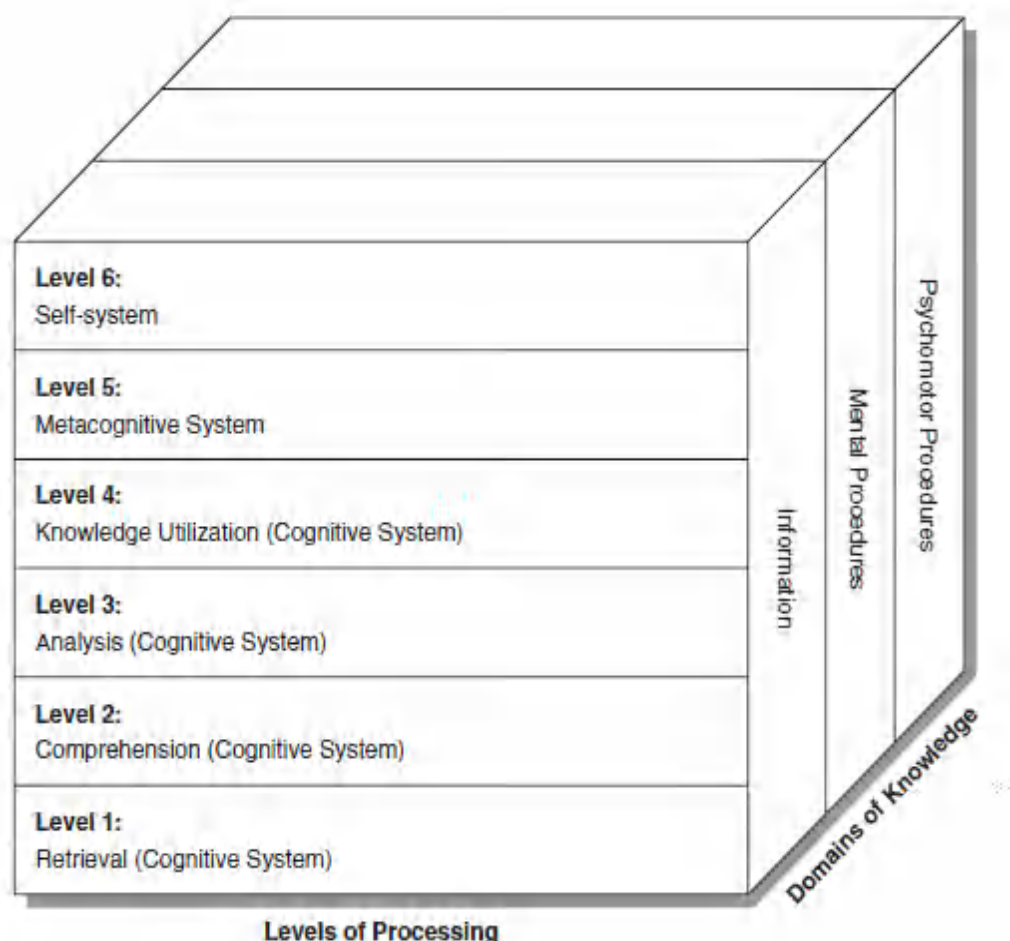


Based on the illustrated model, the new taxonomy of educational objectives was set as follows:

Figure 3.3.

The New Taxonomy

Note. Adopted from Marzano and Kendall (2007, p. 13)



3.5. Achievement Test Planning and Design

3.5.1. Setting the Purpose for Testing

At this level, test designers are guided by a set of questions that need to be answered prior to the item selection phase. Questions like ‘Why are we testing? What should we exactly test? What do we intend to do with the test results? What types of interpretations are needed for these results? , and What for?’ act as a guideline to the construction of a test, which satisfies the purposes set for.

3.5.2. Developing a Test Plan

After setting the purpose for testing, the next step in the process of test planning and design is to decide on a number of test specifications that can ensure a high level of validity to test scores. These specifications, according to Ebel and Frisbie (1991, pp.117-118), must indicate the following:

- 1- Types of test items to be used
- 2- Number of items of each type needed
- 3- Kinds of tasks the items will present
- 4- Number of tasks of each type needed
- 5- Descriptions of content areas to be sampled
- 6- Number of items from each area needed
- 7- Level and distribution of the difficulty of the items.

The specifications listed above serve as a framework or a *plan* for the construction of a test, which provides a high level of results validity. They also help test takers prepare themselves and limit the scope of expectations about how they are going to be tested, about what, and to what extent.

3.5.3. Selection of Item Formats

Different test types can be distinguished on the basis of the test items they encompass. The most common ones are essay, objective and numerical problem tests. Other types include oral and performance tests but they are used in limited contexts compared to the previously stated ones.

3.5.3.1. Essay Tests

Essay test items are those which students have to supply rather than select correct answers (Kubiszyn & Borich, 2013). Two main types of essay tests are discussed in the literature and are, in fact, very familiar among test designers. The first type is *extended response* essays which generally require learners to answer a test question in a lengthy, open-ended

composition. They tend to measure higher-order thinking skills (students' ability to *analyse, synthesise, organise and evaluate*) ; whereas the second type, which is that of *restricted response essays*, is the one in which the learners are supposed to answer a test question in a limited number of lines or words and usually measures lower-order thinking skills (*knowledge, comprehension, and application*).

Ebel and Frisbie (1991, p.123) summarised the points explaining the cases when essay tests, *in general*, are favoured for measuring educational achievement as follows:

- 1- The group to be tested is small, and the test will not be reused.
- 2- The instructor wishes to provide for the development of student skills in written expression.
- 3- The instructor is more interested in exploring students' attitudes than measuring achievement (whether instructors *should* be more interested in attitudes than achievement and whether they should expect an honest expression of attitudes in a test situation seem open to question.)
- 4- The instructor is more confident of his or her proficiency as a crucial essay reader than as an imaginative writer of good objective-test items.
- 5- Time available for test preparation is shorter than time available for test scoring.

Kubiszyn and Borich (2013, p. 160), on the other hand, specified some conditions for which *restricted-response essay items* are best used:

- 1- The instructional objectives require supplying information rather than simply recognizing information. These processes often cannot be measured with objective items.
- 2- Relatively few areas of content need to be tested. If you have 30 students and design a test with six restricted response questions, you will spend a great deal of time

scoring. Use restricted responses when the class size is small, or use them in conjunction with objective items.

- 3- Test security is a consideration. If you are afraid multiple-choice test questions will be passed on or told to other students, it is better to use a restricted response question. In general, a good restricted response essay test takes less time to construct than a good objective test.

3.5.3.2.Objective Tests

The most commonly used item formats in objective tests are true-false, multiple choice, short answer, classification and matching.

3.5.3.2.1. True - false Items

True-False items are very popular among test writers since they are easily and quickly written although good ones take longer time and more effort. According to Ebel and Frisbie (1991, p.133), “The basic reason for using true-false test items is that they provide a simple and direct means of measuring the essential outcomes of formal education.”. They have also summarised the arguments that support the value of true-false items as good measures of educational achievement in four main points:

- 1- The essence of educational achievement is the command of useful verbal knowledge.
- 2- All verbal knowledge can be expressed in propositions.
- 3- A proposition is any sentence that can be said to be true or false.
- 4- The extent of students’ command of a particular area of knowledge is indicated by their success in judging the truth or falsity of propositions related to it.

In order to write good true-false test items, Kubiszyn and Borich (2013, pp. 133-134) gave the following suggestions:

1. The desired method of marking true or false should be clearly explained before students begin the test.
2. Construct statements that are definitely true or definitely false, without additional qualifications. If opinion is used, attribute it to some source.
3. Use relatively short statements and eliminate extraneous material.
4. Keep true and false statements at approximately the same length, and be sure that there are approximately equal numbers of true and false items.
5. Avoid using double-negative statements. They take extra time to decipher and are difficult to interpret.
6. Avoid the following:
 - a. verbal clues, absolutes, and complex sentences.
 - b. broad general statements that are usually not true or false without further qualifications.
 - c. terms denoting indefinite degree (e.g., large, long time, regularly), or absolutes (e.g., never, only, always).
 - d. placing items in a systematic order (e.g., TTFF, TFTF, and so on).
 - e. taking statements directly from the text and presenting them out of context.

3.5.3.2.2. Multiple Choice Items

Another very popular test item type is multiple choice questions generally referred to as MCQs or *selected response* (Bachman, 1996). Kubiszyn and Borich (2013, p.138) state that “Multiple-choice items are unique among objective test items because, contrary to popular opinion, they enable you to measure behavior at the higher levels of the taxonomy of educational objectives”. In their further discussion of this item format, they add that: “Unfortunately, most multiple-choice items are also written at the knowledge level of the taxonomy of educational objectives” (ibid, p.143). This means that the test writer needs to give

special care to setting good objectives that lead to testing higher-level cognitive skills (i.e. comprehension, application, analysis, synthesis or evaluation).

Just like the case of true-false items, the tendency to write good multiple-choice items takes a longer time as this format is probably the most time-consuming type of objective test items to be written. Moreover, Ebel and Frisbie (1991, p.167) consider that “like true-false items, multiple-choice items are developed most conveniently and most appropriately on the basis of ideas expressed or implied in instructional materials”.

Some suggestions for writing well-stated multiple-choice items are summarised by Kubiszyn and Borich (2013, pp.146-147) as follows:

- 1- The stem of the item should clearly formulate a problem. Include as much of the item as possible, keeping the response options as short as possible. However, include only the material needed to make the problem clear and specific. Be concise—don’t add extraneous information.
- 2- Be sure that there is one and only one correct or clearly best answer.
- 3- Be sure wrong answer choices (distractors) are plausible. Eliminate unintentional grammatical clues, and keep the length and form of all the answer choices equal. Rotate the position of the correct answer from item to item randomly.
- 4- In most cases, it is more important for the student to know what a specific item of information is rather than what it is not. So, only use negative questions or statements if the knowledge being tested requires it, and be sure to italicize or somehow highlight the negative(s).
- 5- Include from three to five options (two to four distractors plus one correct answer) to optimize testing for knowledge rather than encouraging guessing. It is not necessary to provide additional distractors for an item simply to maintain the same

number of distractors for each item. This usually leads to poorly constructed distractors that add nothing to test validity and reliability.

- 6- To increase the difficulty of a multiple-choice item, increase the similarity of content among the options.
- 7- Use the option “none of the above” sparingly and only when the keyed answer can be classified unequivocally as right or wrong. Don’t use this option when asking for a better answer.
- 8- Avoid using “all of the above.” It is usually the correct answer and makes the item too easy for students with partial information.

3.5.3.2.3. Short-answer Items

Ebel and Frisbie (1991, p.179) argue that “a short-answer test item aims to test knowledge by asking examinees to supply a word, phrase, or number that answers a question or completes a sentence. Completion and fill-in-the-blank are other common labels for short-answer items”. On that basis, *completion items* in Kubiszyn and Borich's (2013) classification of objective tests’ item formats refers to short-answer items. The same type of test items is sometimes referred to as *limited production response* (Bachman, 1996).

Some characteristics of short-answer items that have been explained by Ebel and Frisbie (1991) can be summarised as follows:

- What short-answer items can test (*factual information*) is much more limited than what true-false or multiple-choice items can test,
- short-answer items are much less affected by guessing,
- they are supposed to test recall rather than recognition, which makes them more demanding and more valid to test learners’ achievement,
- They are quite easy to write and also efficient as they provide a number of separate scorable sections which makes it less burdensome for the learner,

- They are the best item type option when students have reading or writing problems.

Some suggestions for writing short-answer items were presented by Kubiszyn and Borich (2013, p.150) as follows:

- 1- If at all possible, items should require a single-word answer or a brief and definite statement. Avoid statements that are so indefinite that they may be logically answered by several terms.

a. Poor item: World War II ended in _____.

b. Better item: World War II ended in the year _____.

- 2- Be sure the question or statement poses a problem to the examinee. A direct question is often more desirable than an incomplete statement (it provides more structure).
- 3- Be sure the answer that the student is required to produce is factually correct. Be sure the language used in the question is precise and accurate in relation to the subject matter area being tested.
- 4- Omit only keywords; don't eliminate so many elements that the sense of the content is impaired.

a. Poor item: The _____ type of test item is usually more _____ than the _____ type.

b. Better item: The completion type of test item is usually graded less objectively than the _____ type.

- 5- Word the statement such that the blank is near the end of the sentence rather than near the beginning. This will prevent awkward sentences.
- 6- If the problem requires a numerical answer, indicate the units in which it is to be expressed.

3.5.3.2.4. Matching Items

Matching items are very popular like true-false and MCQs but to write good ones, special attention should be given to some details that may not be that obvious for novice test writers. For Ebel and Frisbie (1991):

Matching test items occur in clusters composed of a list of premises, a list of responses, and directions for matching the two. In many clusters, the distinction between premises and responses is simply in the names given to them. The two lists can be interchanged without difficulty. (pp. 182-183)

Some suggestions Ebel and Frisbie (1991, pp. 184-185) give for writing good matching test items are summarised as follows:

- 1- Choose homogeneous premises and responses for any matching cluster,
- 2- Make the lists of promises and responses relatively short,
- 3- Do not attempt “perfect” matching, in which each response is matched once and only once to each of the premises,
- 4- Provide directions that clearly explain the intended basis for matching,
- 5- Arranging responses or premises or both in alphabetical order usually prevents give-away clues that can occur in item writing,
- 6- If the responses are numerical quantities, arrange them in order from low to high,
- 7- Use the longer phrases as premises, the shorter as responses.

3.5.4. Quantity of Test Items

The quantity or number of test items to be included in a test is largely determined by the amount of time allotted to it in addition to other factors including the degree of difficulty of test items, their nature (essay, true/false, MCQs, ...etc.), as well as the learners’ level of proficiency and the educational objective set before the test. Most tests and examinations take no more than

3 hours to be done depending on the subject being examined and the educational level (primary, middle, or secondary school).

Giving importance to the number of items to be included in a test is simply due to the ultimate goal of giving all the students, who went through the same course or instructional unit, to have equal chances to answer most if not all the test items within the time provided at their normal rates. Test anxiety is another reason for which test designers need to make a balanced test in terms of the number of items (taking into account the above-mentioned factors) and the time allotted. This can be explained by the fact that the more test takers feel they can answer all the test items within the time limits, the more accurate they would be in giving answers and vice versa.

Another point to consider when thinking about the number of items to be included in a test is that the test scores are more reliable when the number of test items is higher as the test will cover more elements and the test items (*sample*) are going to be more representative of the course content or the instructional unit (*population*) (Ebel & Frisbie, 1991).

3.5.5. Distribution of Test Items' Difficulty Level

Items of a particular test are selected based on their level of difficulty and in terms of their ability to distinguish between students vis-à-vis their mastery of the instructional content. Hence, the required level of performance for the different levels of success in the test should be specified (Hughes, 1989).

Some tests tend to end up with examinees having homogeneous scores. This is due to the fact that the test includes items that any student who has studied successfully can answer. This kind of test cannot distinguish between students of different levels of performance (poor, weak, average, good, and excellent).

Many norm-referenced tests tend to include items that only better learners can answer in addition to some items that all examinees who have attended the same instructional content can respond to. Such kinds of tests can make the distinction between learners by establishing a

rank order of learners' scores as well as a higher standard deviation compared to the previous test type.

3.6. Describing and Summarising Achievement Test Results

3.6.1. Summarising Data

3.6.1.1. Frequency Distribution

3.6.1.1.1. The List

The simplest way to organise and make some sense of order to students' scores on a test is to list them in ascending or descending order. However, this method allows the interpretation of the scores of a small group of students (e.g. one single class) only and as the number of scores increases, the usefulness of such a method decreases. Below is an example of 20 students' class scores arranged in ascending order:

6.5	12	14.5	17
9.5	12.75	15	17.25
9.5	13	15.75	18
9.75	14	16.25	18.5
10	14.25	16.5	19.25

A short list of scores like the one above can give us an idea about the highest scores, the lowest scores and the middle scores knowing that the test is scored out of 20 points. Nevertheless, how would the list and its interpretation be if the number of scores was 100, 2,000 or even more (e.g. in the case of large-scale examinations)? This will require other methods of organising and summarizing data. The next point deals with *Simple Frequency Distribution*, which is another method for tabulating data that considers all scores including those that are missing.

3.6.1.1.2. Simple Frequency Distribution

Simple Frequency Distribution does not seem to be of great help when it comes to classifying scores for which the amount of variation is high. An example of this is shown in Table 3.1., which shows the *Simple Frequency Distribution* of the 20 students' class scores illustrated in

the point of *the list* previously discussed. Similarly, the interpretation of the results does not seem to be an easy task and the high number of scores and variations makes it confusing and even difficult to make sense of the data provided especially since there are a lot of missing scores.

Table 3.1.

Simple Frequency Distribution of 20 Students' Class Scores

Score	Frequency	Score	Frequency
6.5	1	13	1
6.75	0	13.25	0
7	0	13.5	0
7.25	0	13.75	0
7.5	0	14	1
7.75	0	14.25	1
8	0	14.5	1
8.25	0	14.75	0
8.5	0	15	1
8.75	0	15.25	0
9	0	15.5	0
9.25	0	15.75	1
9.5	2	16	0
9.75	1	16.25	1
10	1	16.5	1
10.25	0	16.75	0
10.5	0	17	1
10.75	0	17.25	1
11	0	17.5	0
11.25	0	17.75	0
11.5	0	18	1
11.75	0	18.25	0
12	1	18.5	1
12.25	0	18.75	0
12.5	0	19	0
12.75	1	19.25	1

3.6.1.1.3. Grouped Frequency Distribution

One of the variations of *Simple Frequency Distribution* is *Grouped Frequency Distribution*, which tends to be a more practical and goal-achieving method for summarizing scores in a way to make sense of the obtained data. Contrary to *Simple Frequency Distribution*, where each

possible score is taken as a category, in *Grouped Frequency Distribution*, intervals or ranges of scores are considered as categories. An example of *Grouped Frequency Distribution* is the one shown in Table 2.4. illustrating the same 20 scores previously mentioned.

Table 3.2.

Grouped Frequency Distribution of a 20 Students' Class Scores

<i>Interval</i>	<i>f</i>
18-21	3
15-18	6
12-15	6
9-12	4
6-9	1

There are five (5) steps to follow in order to construct a grouped frequency distribution:

- 1- Determine the range of scores (R): $R = H - L$, where (H) refers to the highest score and (L) to the lowest. e.g. $R = 19.25 - 6.5 = 12.75$
- 2- Determine the number of intervals, which is an arbitrary decision. It can be 5, 8, 10, 12...etc. depending on the number of scores and the frequencies in each interval.
- 3- Divide the range by the number of intervals selected and round the result to the nearest odd number. This is known as the interval width symbolised by i .

$$i = \frac{R}{\text{number of intervals}} \quad \text{e.g. } i = \frac{12.75}{5} = 2.55 \text{ (rounded to the nearest odd number} \quad = 3)$$

- 4- Make the interval column by ensuring that the lower limit is a multiple of the interval width and the higher limit is one point lower than the lower limit of the next interval. In our case, we have scores that are in the form of decimal numbers. Therefore, the higher limit of each interval is the same as the lower limit of the next one. This is how the interval column would look like:

Lower limit	Higher limit
21	24
18	21
15	18
12	15
9	12
6	9
3	6

Hence, we retain only the intervals that our scores fall within and exclude the other ones. The list of intervals in our example is as follows:

Intervals

18-21

15-18

12-15

9-12

6-9

- 5- Construct the frequency (f) column by giving the number of scores captured in each interval. The number of scores of the exact value of the higher limit in each interval is not counted in the frequencies of that interval.

The result is a table like Table 3.2., which includes the intervals and the frequencies of scores captured in each interval.

Among the advantages of the grouped frequency distribution of scores is that it compresses the size of the table and makes the data much easier to interpret. However, it has the drawback of having information about individual scores lost (Kubiszyn & Borich, 2013).

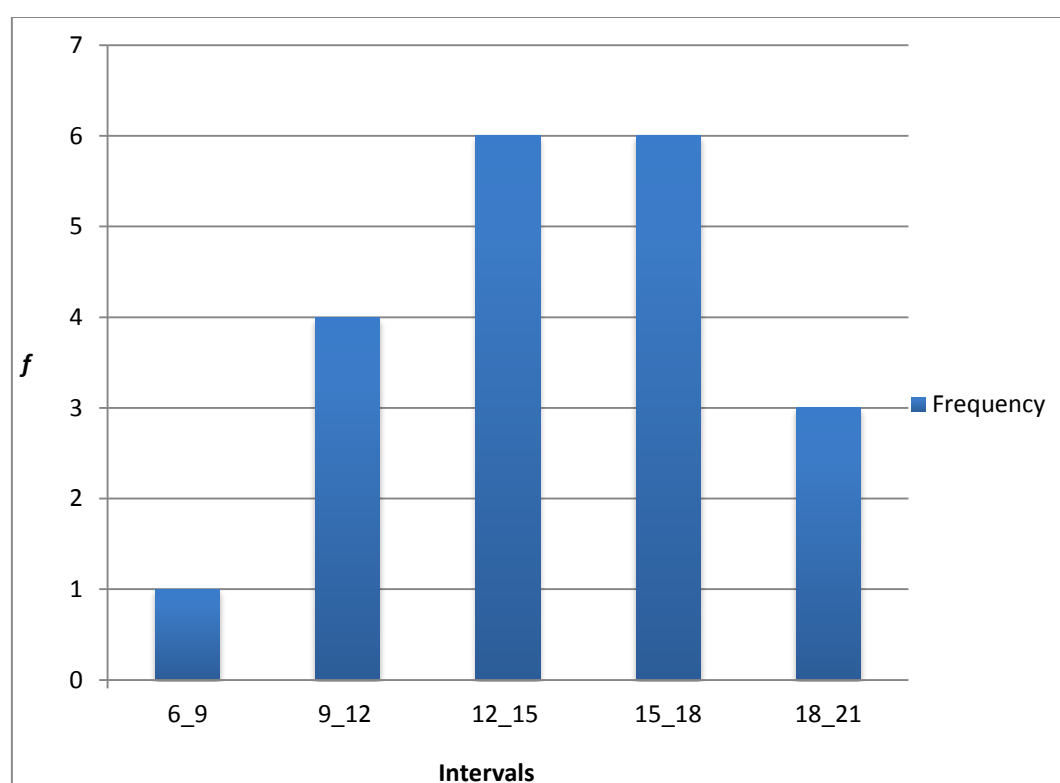
3.6.1.2. Graphing Data

Graphs are very important in the process of presenting and interpreting test results. Cohen & Swerdlik (2009) state that “One picture is worth a thousand words and one purpose of representing data in graphic form is to convey information at a glance” (p.95). Accordingly, three main types of graphic representation of data are found in literature: *the graph bar* (or *histogram*), *the frequency polygon* (or *line graph*) and *the smooth curve*.

Figure 3.4., is a *histogram* presenting the data of the grouped frequency distribution for the 20 students’ scores previously displayed. As it can be noticed, the interpretation of the results can be done without much effort as the representation of data is straightforward i.e. the higher the bar, the greater the frequency of scores falling into that interval.

Figure 3.4.

Histogram Based on a Grouped Frequency Distribution

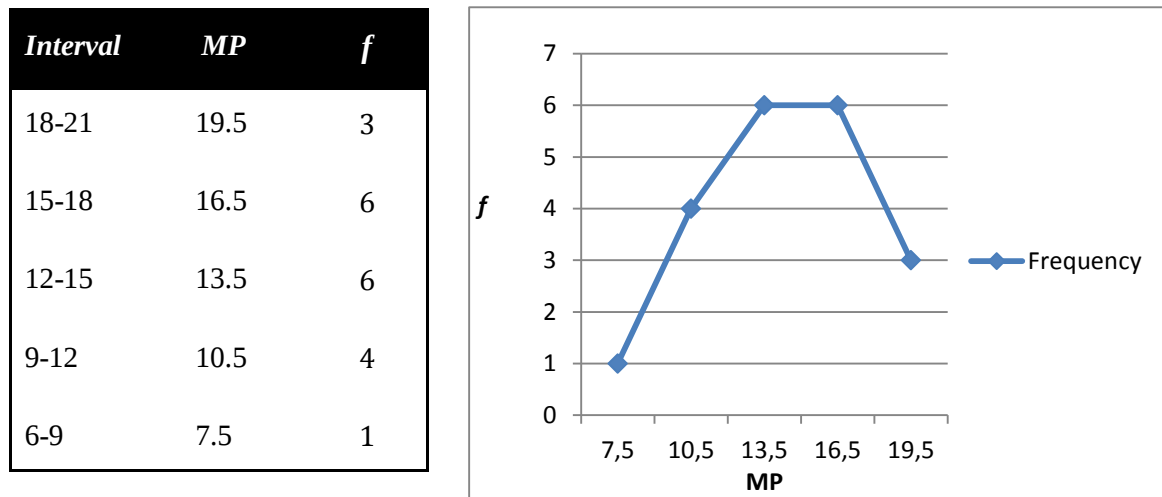


A *frequency polygon* can also be used to give a picture of the same test scores. Just like a histogram, the horizontal line is where the scores are depicted while the vertical line is that of

the frequency scale. The difference between the two graphical methods is that the frequency polygon “uses straight lines to connect the midpoint (MP) of each interval rather than bars or columns to show the frequency with which scores occur.” (Kubiszyn & Borich, 2013, pp. 271-272). This is shown in Figure 3.5, which uses the scores and the grouped frequency distribution of the same group of students:

Figure 3.5.

Grouped Frequency Distribution with a Frequency Polygon

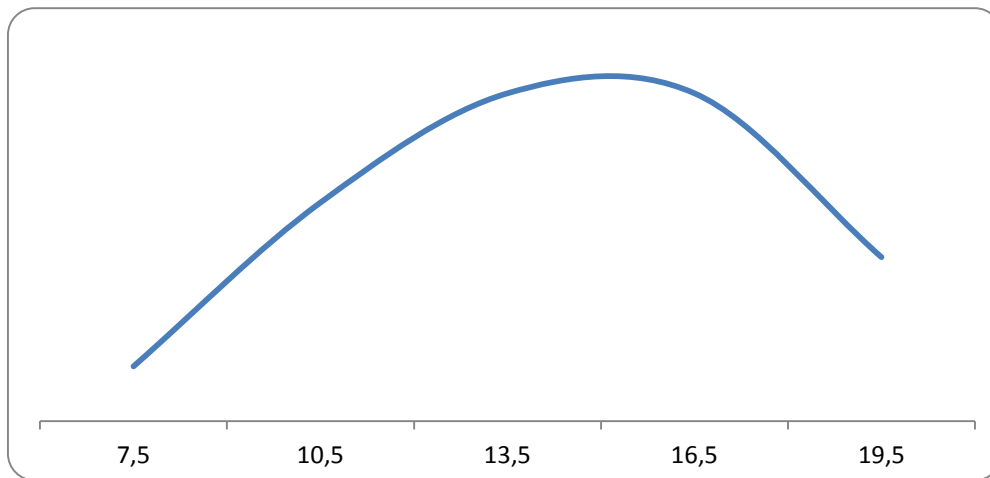


Instead of the frequency polygon illustrated in Figure 3.5., a *smooth curve* (Figure 3.6.) can also serve as a means to depict the shape of the distribution of scores rather than representing any set of data precisely. The main differences between *a frequency polygon* and *a smooth curve* are in the latter:

- The Y axis (vertical line) used to present the frequencies is deleted;
- Curved lines are used instead of straight ones.

Figure 3.6

A Smooth Curve



There are two major types of smooth curves: *symmetric* and *asymmetric (or nonsymmetric)*. To make the distinction between them, Ebel & Frisbie (1991) explain that “a symmetric curve is one that has two halves, each a mirror image of the other...Nonsymmetric or asymmetric curves are those that do not have this property. Such curves are said to be *skewed*.” (p.58). *Skewed curves* are divided into two categories: *positively skewed* and *negatively skewed curves*. Kubiszyn and Borich (2013) point that:

A positively skewed distribution results from an asymmetrical distribution of scores. In this case, the majority of scores fall below the middle of the score distribution. There are many low scores but few high scores... A negatively skewed distribution also results from an asymmetrical score distribution. In this type of distribution, the majority of scores fall above the middle of the score distribution. There are many high scores but few low scores. (pp. 276-277)

Hence, the smooth curve illustrated in Figure 3.6 is *asymmetric* and to be more precise it is *negatively skewed*.

3.6.2. Describing and Summarising Score Distributions

3.6.2.1. Measures of Central Tendency

Cohen and Swerdlik (2009) define *a measure of central tendency* as “...a statistic that indicates the average or midmost score between the extreme scores in a distribution” (p.80). Three major measures of central tendency are commonly used: *the mean*, *the median* and *the mode*. These are reviewed separately in this section with some reference to the formulas used to calculate them.

3.6.2.1.1. (Arithmetic) Mean

Perhaps, this is the most commonly used measure of central tendency. In everyday use, it is generally referred to as *the average*. In the book by OpenStax College (2015), it is pointed out that:

When each value in the data set is not unique, the mean can be calculated by multiplying each distinct value by its frequency and then dividing the sum by the total number of data values. The letter used to represent the sample mean is an x with a bar over it (pronounced “x bar”): $\bar{x}$. (p.70)

It is also possible to find some references using the symbol M to refer to the mean score (Kubiszyn & Borich, 2013). According to the above explanation, the mean score can be calculated using the following formula:

Formula

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

plain English version

$$\text{Mean} = \frac{\text{sum of all the scores}}{\text{total number of scores}}$$

3.6.2.1.2. Median

This is the second most frequently used measure of central tendency after the average (arithmetic mean). Ebel and Frisbie (1991) define it as “the score above which and below which exactly half of the scores are found: the middlemost score. For the scores 5,4,3,2, 1, the median

is 3; for the scores 9, 7, 5, 2, the median is 6, halfway between 7 and 5.” (p.59). The main property of the median is the fact that it is not influenced by extreme scores since only the middle or two middle scores are considered in determining it (Kubiszyn & Borich, 2013).

Therefore, the median (abbreviated MDN) of the 20 students’ scores provided as an example in the previous sections would be calculated as follows:

6.5	12	14.5	17
9.5	12.75	15	17.25
9.5	13	15.75	18
9.75	14	16.25	18.5
10	14.25	16.5	19.25

$$\text{MDN} = \text{Average of the two middle scores } i.e. \text{ MDN} = \frac{14.25 + 14.5}{2} = 14.34$$

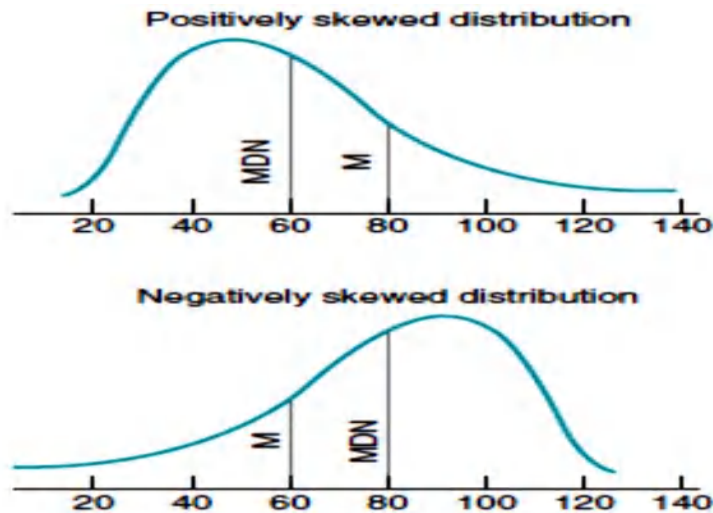
To show the difference between the *mean* and the *median* as well as to stress the importance of this later in the case of *skewed distributions*, Kubiszyn and Borich (2013) point that:

Since the median is not affected by extreme scores, it represents central tendency better than the mean when distributions are skewed. In skewed distributions, the mean is pulled toward the extremes, so that in some cases it may give a falsely high or falsely low estimate of central tendency. Since the median, on the other hand, is the score above and below which half the scores lie, it always indicates the center of the distribution. (p.282)

Figure 3.7.

Importance of the Median to Represent Central Tendency in Skewed Distributions

Note. Adopted from Kubiszyn and Borich (2013, p.282)



3.6.2.1.3. Mode

The *mode* is generally defined as the most occurring value. However, it is not frequently used as a measure of central tendency as it is the least stable one. Cohen and Swerdlik (2009) make a very important distinction between the three measures of central tendency by saying:

Because the mode is not calculated in a true sense, it is a nominal statistic and cannot legitimately be used in further calculations. The median is a statistic that takes into account the order of scores and is itself ordinal in nature. The mean, an interval-level statistic, is generally the most stable and useful measure of central tendency. (pp. 86-87)

In a *normal distribution*, the *mean*, the *median* and the *mode* all have the same value. In a *positively skewed* distribution, the mode has the lowest value, the mean the highest of all and the median is the intermediate value. In a *negatively skewed* distribution, on the other hand, it is quite the opposite (i.e. the mean=the lowest / the mode= the highest value).

3.6.2.2.Measures of Variability

Variability is a statistical term that refers basically to the “spread out” of a group of scores (OpenStax College, 2015). *Measures of variability* are, then, statistics that describe the amount of variation in a distribution (Cohen & Swerdlik, 2009). The main measures of variation include *the range, the interquartile range, the semi-interquartile range, the variance and the standard deviation.*

3.6.2.2.1. Range

This is the simplest measure of variability to be calculated and one that has already been used in this chapter to calculate the interval. The range is simply determined by subtracting the lowest score from the highest score. In the example of the 20 students’ scores, the range (R) is calculated as follows:

$$R = H - L \text{ (H refers to the highest score / L refers to the lowest score)}$$

$$\text{Therefore: } R = 19.25 - 6.5 = 12.75$$

However, Cohen and Swerdlik (2009, p.87) criticized this measure of variability by saying that:

As a descriptive statistic of variation, the range provides a quick but gross description of the spread of scores. When its value is based on extreme scores in a distribution, the resulting description of variation may be understated or overstated.

Consequently, other measures of variability would be more useful and practical.

3.6.2.2.2. Interquartile Range and Semi-interquartile Range

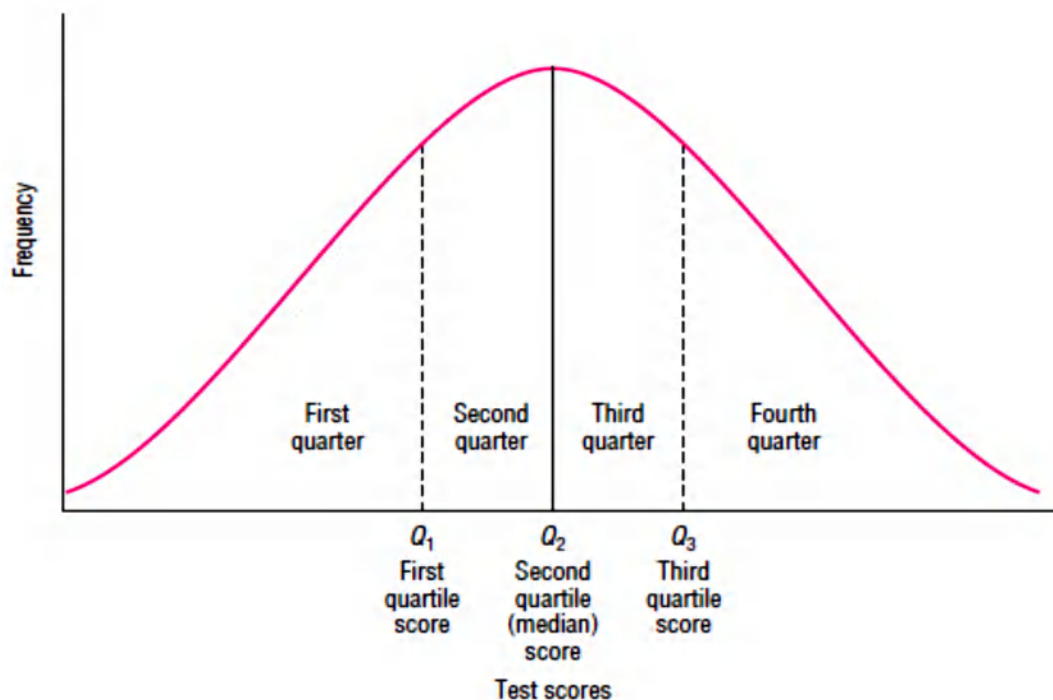
Before discussing these two notions, it seems quite necessary to explain the term *quartile* as it appears in both terms. *Quartiles* are numbers that separate data into quarters. It is worth mentioning here that *a quartile* refers to a specific point whereas *a quarter* refers to an interval. In any distribution, three quartiles can be identified as Q1, Q2 and Q3 respectively. Q2 in a distribution of scores is the same as *the median*, which means that it is the midpoint of the

distribution. As a result, Q_1 is the middle value of the lower half of the data, and Q_3 is the middle value of the upper half (See figure 3.8).

Figure 3.8.

Quartered Distribution

Note. Adopted from Cohen and Swerdlik (2009, p.100)



The *interquartile range (IQR)* is, by definition, “a number that indicates the spread of the middle half or the middle 50% of the data. It is the difference between the third quartile (Q_3) and the first quartile (Q_1). $IQR = Q_3 - Q_1$ ” (OpenStax College, 2015, p. 63). The *semi-interquartile range (SIQR)* is simply defined as the IQR divided by 2. One advantage of IQR and SIQR is reported by Cohen and Swerdlik (2009, p.88) as follows:

Knowledge of the relative distances of Q_1 and Q_3 from Q_2 (the median) provides the seasoned test interpreter with immediate information as to the shape of the distribution of scores. In a perfectly symmetrical distribution, Q_1 and Q_3 will be exactly the same distance

from the median. If these distances are unequal then there is a lack of symmetry. This lack of symmetry is referred to as skewness.

Kubiszyn and Borich (2013) further argue that:

Although the SIQR has the advantage of not being influenced by extreme scores, it has the disadvantage of being determined ultimately by only half the scores in a distribution. Any time all scores in a distribution do not enter the computation of a statistic, an element of error is introduced. (p.292)

3.6.2.2.3. Variance

Variance is another method to measure variability but in terms of how close the scores in the distribution are to the middle of the distribution (Lane, n.d.). He further explains that “Using the mean as the measure of the middle of the distribution, the variance is defined as the average squared difference of the scores from the mean.” (ibid, p.147).

The formula for calculating the variance in a population is, therefore, put like this:

$$\sigma^2 = \frac{\Sigma(X - \mu)^2}{N}$$

Where: σ^2 is the population *variance*, μ is the *mean*, and N is the *number of scores*.

On the other hand, if the variance in a sample is calculated to determine the variance in a population, the formula would be as follows:

$$s^2 = \frac{\Sigma(\bar{X} - x)^2}{N - 1}$$

Where: s^2 is the sample *variance*, $\bar{X}$ is the *sample mean*. N-1 is the *sample size -1*

3.6.2.2.4. Standard Deviation

The *standard deviation* can be defined simply as a measure of variability equal to the square root of *the variance*. It is considered the most accurate and a very useful measure of variability since it considers all scores in a distribution in its computation (Kubiszyn & Borich, 2013). A number of symbols are used to refer to the standard deviation like *s*, *S* or *SD* in addition to the lowercase Greek letter sigma (σ). According to the stated definition, the formula for the population standard deviation is as follows:

$$\sigma = \sqrt{\frac{\Sigma(X - \mu)^2}{N}}$$

Whereas, to calculate the sample standard deviation, this formula is used:

$$s = \sqrt{\frac{\Sigma(X - x)^2}{N - 1}}$$

The importance of the standard deviation, as Kubiszyn and Borich (2013) put it, lies in the fact that “it stands to reason that large SD values indicate greater score variability than do small SD values” (p.298).

3.6.2.3.Normal Distribution

This is the most widely used and also the most important distribution in the field of statistics. It is also referred to as the ‘bell curve’, the ‘Gaussian curve’ (Lane, n.d.), and the ‘normal curve’ (Cohen & Swerdlik, 2009). It is defined by Ebel and Frisbie (1991) as:

A frequency polygon that is symmetric, unimodal, and mesokurtic is, in plain talk, bell-shaped. It is also called a *normal distribution*. The normal curve is a sort of standard that is used frequently as a reference point for describing a set of test scores. For example, we might

make statements like: 'it's nearly normal.' or 'it's a bit flatter than the normal curve.' or 'it's like a normal curve with fat tails'. (p.62)

The normal curve is, in fact, hypothetical as no distribution matches exactly the normal distribution. This latter has some fixed properties summarised by Lane (n.d., pp.250-251) as follows:

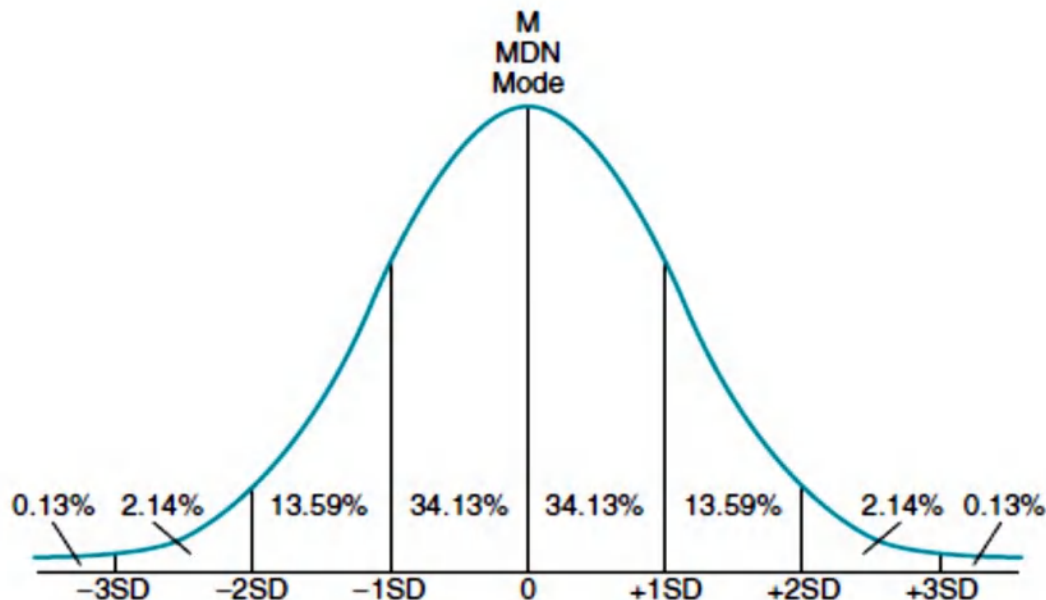
1. Normal distributions are symmetric around their mean.
2. The mean, median, and mode of a normal distribution are equal.
3. The area under the normal curve is equal to 1.0.
4. Normal distributions are denser in the centre and less dense in the tails.
5. Normal distributions are defined by two parameters, the mean (μ) and the standard deviation (σ).
6. 68 % of the area of a normal distribution is within one standard deviation of the mean.
7. Approximately 95% of the area of a normal distribution is within two standard deviations of the mean.

An illustration of a normal curve and the area below it is presented in Figure 3.9:

Figure 3.9.

Percent of Cases under Portions of the Normal Curve

Note. Adopted from Kubiszyn and Borich (2013, p.300)



3.6.2.4. Standard Scores

Standard scores are simply raw scores that have been converted from one scale to another in order to make them easily interpretable. In statistics books, two main systems for calculating standard scores are found and will be referred to later in Chapter Five. These are the *z scores* and the *t scores*.

3.6.2.4.1. *z* Scores

A *z* score, according to Cohen and Swerdlik (2009):

...results from the conversion of a raw score into a number indicating how many standard deviation units the raw score is below or above the mean of the distribution... In essence, a *z* score is equal to the difference between a particular raw score and the mean divided by the standard deviation. (p.96)

The last part of this quotation can be summarized by the following formula:

$z = \frac{X - \mu}{\sigma}$, where z is the value on the standard normal distribution, X is the value on the original distribution, μ is the mean of the original distribution, and σ is the standard deviation of the original distribution.

Since the standard normal distribution has a mean of 0 and a standard deviation of 1, the z score is equal to the number of standard deviations below (or above) the mean. For example, a z of -1.25 represents a value of -1.25 standard deviations below the mean.

3.6.2.4.2. t Scores

Here, contrary to the z score, the scale has a mean is set at 50 and a standard deviation set at 10. This means that this standard score system's scale ranges between five standard deviations below the mean and the same number of standard deviations above the mean. The difference between this system and the previous one is the absence of negative values. Hence, t scores range from 0 to 100.

Conclusion

Achievement tests are among the most widely used test types, especially in educational systems. The emphasis that is given to this test type in this chapter is due to the nature and topic of this research. The fact that the BEM and BAC exams are final tests that measure learners' achievements at the end of a learning cycle (4 years and 3 years respectively) and determine whether examinees can move to the next educational level, it was more than necessary to deal with all the elements discussed in the chapter in order to proceed with the analysis of the exam papers (Chapter Four/Section Two) as well as comparing male and female examinees' performances in these exams (Chapter Five). The content of this chapter represents a solid foundation and a comprehensive background on which to build in order to arrive at reliable and meaningful results in the coming chapters.

Chapter Four

Thematic and Task-Based Analysis of English Language Textbooks and BEM and BAC

Exam Papers

Table of Contents

Introduction	112
Section One: Thematic and Task-Based Analysis of ‘On the Move’ and ‘New Prospects’ ..	112
4.1.1. Presentation of the Textbooks	112
4.1.1.1. On the Move (OTM)	112
4.1.1.2. New Prospects	115
4.1.2. Content-based Analysis	117
4.1.2.1. Topics / Themes	118
4.1.2.2. Activity/Task-Based Evaluation of “On the Move” and “New Prospects”	138
Section Two: Thematic and Items’ Analysis of English Language BEM and BAC Exam Papers	142
4.2.1. The BEM English Language Exam	142
4.2.1.1. Description of the Exam Paper	142
4.2.1.2. Analysis of the Exam Papers	143
4.2.2. The Baccalaureate English Language Exam	146
4.2.2.1. Description of the Exam Papers	146
4.2.2.2. Analysis of the Exam Paper	148
4.2.3. Summary of the Findings	154
Conclusion	155

Introduction

In Chapter two, a theoretical discussion of materials design and evaluation was presented. Two major elements that any analysis of materials or textbook evaluation can deal with are discussing the topics around which the units of the textbook are built and also identifying the task and activity types that are used to sustain and boost the students' learning aiming to master the linguistic, communicative, and interactional content of the textbook. This is exactly the focus of *section one* of this chapter, in which the textbooks used in the final years of both middle and secondary school education 'On the Move' and 'New Prospects' are analysed. This is not done as a primary aim on its own as the analysis results will be used in the later chapters together with the exam papers analysis, which is done in *section two*, to reflect on them and draw conclusions on the extent of the match/mismatch between their contents.

The selected exam samples were BEM 2014 and BEM 2019 in addition to BAC 2014 and BAC 2019 (Scientific and Foreign Languages streams). This selection was not arbitrary, but rather was done based on some specifications regarding the examinees' populations (See Chapter Five) and also the overall learning/teaching conditions that characterized both school years, i.e. 2013/2014 and 2018/2019. Yet, before dealing with the analysis, it is necessary to make a general presentation of both textbooks and also the exam papers in their respective sections.

Section One: Thematic and Task-Based Analysis of 'On the Move' and 'New Prospects'

4.1.1. Presentation of the Textbooks

4.1.1.1. On the Move (OTM)

It was first published in 2006 by the Official Authority for School publication on behalf of the Ministry of National Education in Algeria to be used by fourth-year middle school pupils throughout the country. S.A. Arab and B. Riche are the main authors of this textbook. The

version of the textbook being analysed and evaluated in this research is the one published for the 2018-2019 school year and it is its last edition of this, which was later replaced by My Book of English middle school year 4 in the year after. “On the Move”, as its authors put it, “takes up from Spotlight on English, Book Three while developing features of its own” (p.8). The authors, in the section entitled ‘To the Teacher’, summarise these features in eight (8) points as follows:

- 1- The clear-cut distinction between the receptive phase and the production phase;
- 2- The inclusion of an objective, graded end-of-the-file evaluation (Progress Check);
- 3- The streamlining of the cultural component. This latter is considered in-built;
- 4- The strengthening of the correlation between the primary skills, the social skills and the project;
- 5- The slotting-in of a soft storyline involving a number of recurrent characters throughout the book;
- 6- The foregrounding of grammar, both in theory and in practice, notably through checking and cross-referencing;
- 7- The widening of the intercultural scope to new horizons (USA, India, and Australia) through comparison and contrast with Algeria;
- 8- And finally, the development of student autonomy through “survival strategies” (coping) and research tasks involving group work and peer evaluation.

The textbook is intended, as its authors put it, to provide the basic language learning to meet the requirements of the brevet examination while creating a joyful and funny learning environment by including songs and games within the files.

“On the Move” consists of 192 pages and comprises six (6) files organised as follows:

- 1- **File One:** It’s My Treat (pp.16-40);
- 2- **File Two:** You Can Do It (pp.41-64);
- 3- **File Three:** Great Expectations(pp.65-89);
- 4- **File Four:** Then and Now (pp.90-116);
- 5- **File Five:** Dreams, Dreams... (pp.117-140);
- 6- **File Six:** Fact and Fiction (pp.141-164).

Each one of the abovementioned files is divided into two main sections: Language Learning and Skills Building respectively. The first Section in each file i.e. Language Learning includes three basic lessons, which are:

- Listen and Consider;
- Read and Consider;
- Words and Sounds.

A “Take a break” page separates the two sections of the file. It contains either a song or everyday English content allowing the learner to enjoy a non-traditional learning environment; hence, practising their English with fun. Similarly, the Skills Building section is composed of three main lessons:

- Research and report;
- Listening and Speaking;
- Reading and Writing.

All six files conclude with a “**Project round-up**”, where students are asked to work in small groups on a project outside the classroom. Here students need to make use of all the topics, knowledge, and language items, expressions, and structures learnt within the file in addition to conducting some research and involving other skills (drawing, design...etc.) to accomplish the task. Finally, students are allowed to check their progress in a section entitled

“**Where do we stand now?**” and also to have some fun at the end of the file and explore the British/American culture in “**Time for**”.

The coursebook ends up with:

- Scripts (pp. 165-175);
- Grammar reference (pp.176-190);
- Phonetic symbols (p.191);
- Acknowledgement (p.192).

4.1.1.2.New Prospects

The changes that took place in the educational curricula and teaching/learning materials in Algeria starting from 2016 with the implementation of the second-generation textbooks in the middle school did not affect the secondary school level. '*At the Crossroads*' (1st Year), '*Getting Through*' (2nd Year) and '*New Prospects*' (3rd Year) were used starting from 2005 and adapted to fit the second-generation curriculum.

B. Riche and S.A. Arab are the main authors of this series of textbooks. However, other people participated in the design of only parts of it. Those are:

- H. Hami and H. Ameziane (*At the Crossroads* SE1 & *Getting Through* SE2)
- M. Bensemmane (*Getting Through* SE2 & *New Prospects* SE3)
- K. Louadj (*At the Crossroads* SE1)

Unlike the second-generation English language textbooks (*My Book of English*) used in middle schools and published by Casbah Editions, which is a private publisher, the three textbooks used by Algerian secondary school learners are published by the National Authority for School Publications. This latter is the governmental and official publishing authority for all the national education materials in Algeria. A more detailed description of the textbooks is presented in what follows.

'*New Prospects*' is the third book in the series of textbooks designed for secondary school students in Algeria. It builds upon the knowledge and skills learnt through the use of '*At the Crossroads*' and '*Getting Through*' by trying to cover learners' interests taking into consideration the different streams they belong to.

The authors of this textbook consider that '*New Prospects*' "complies with the new English syllabus for SE3 as laid out by the National Curriculum Committee of the Ministry of National Education in March 2006." (p.4). In the same context, the textbook designers precise that three main features of the syllabus were given careful consideration in designing this textbook:

- The fact that the *Baccalaureate* is exclusively of the written mode;
- The emphasis on a thematic orientation;
- The need to cater for the pedagogical requirements of all **Baccalaureate** streams.

The textbook, according to its authors again, is designed based on the competency-based, learner-centred, project-gearred approach.

New Prospects contains six (6) units with four mandatory units for each stream. The total number of pages is 271 taking into account the Listening scripts, Grammar reference, Resources portfolio, and the Acknowledgements found at the end of the textbook. The units are organized as follows:

- **Unit One:** Exploring the Past (pp.14-44)
- **Unit Two:** Ill-Gotten Gains Never Prosper (pp.45-73)
- **Unit Three:** Schools: Different and Alike (pp.74-105)
- **Unit Four:** Safety First (pp.106-134)
- **Unit Five:** Are We Alone? (pp.135-164)
- **Unit Six:** We are a Family! (pp.165-195)

Third-year secondary school students from different streams deal with four units only out of six depending on the units' contents and the themes students are supposed to discuss based on the stream they follow. Hence, the SE3 syllabus provides the following distribution:

- Common Streams (3 hours/week): Units 2, 4, 5, & 6.
- Literature and Philosophy & Foreign Languages (4 hours/week): Units 1, 2, 3, and 6.

Each of these units contains two main parts, which are: **Language Outcomes** and **Skills and Strategies Outcomes** separated by a 'Take a break' section. Each part is divided into two sequences: **Language Outcomes** is divided into *Listen and Consider* and *Read and Consider*; whereas, the **Skills and Strategies Outcomes** part includes two other sequences, which are *Listening and Speaking* and *Reading and Writing*. Moreover, this part includes two additional sections entitled *Research and report* and *Project outcomes* respectively. These are aimed at developing students' technological, research and report skills and strategies. The units end up with an **assessment section** that consists of two parts: *Language assessment*, which is in the form of a grid with a scale, and *Skills and strategies assessment*, which instructs the learners to take a test assigned to them by their teacher on one of the texts in the **Resources Portfolio** at the end of the textbook. Finally, the unit concludes with either a song or a poem on one page entitled *Time for...* giving the students some time for fun before starting a new unit.

4.1.2. Content-based Analysis

To achieve the aims of the study, and because the learners' results analysed in the next chapter are those achieved in the BEM and BAC exams, the content-based analysis will be limited to two (2) textbooks, which are "On the Move", and "New Prospects". The reasons behind this are:

- In both levels, the textbooks are designed in the form of a series with a head of the project ensuring continuity in the building up of skills and competencies in addition to a gradation of abilities and strategies necessary to deal with the activities and tasks from

one level to another. The intended final objectives of the series are supposed to be present in and to be achieved within and by the use of the last textbook of each.

- The topics of the BEM and BAC exams are limited to the ones discussed in the files/sequences of the fourth-year middle school and third-year secondary school textbooks respectively;
- The analysis of learners' achievements in BEM 2014 and 2019 exams, in the next chapter, made it necessary to analyse and evaluate OTM bearing in mind that it was adapted to the second-generation curriculum and used in the 2018/2019 school year instead of MBOE 4. This latter started to be used in the following year.

4.1.2.1.Topics / Themes

4.1.2.1.1. On the Move (OTM)

OTM, based on the syllabus presented at the beginning of the textbook, discusses within the six units it includes the following topics:

- **File 01 (It's My Treat):** Food and Drink
- **File 02 (You Can do It):** Citizenship- Sustainable Development
- **File 03 (Great Expectations):** People and Places
- **File 04 (Then and Now):** Customs and Mores
- **File 05 (Dreams, Dreams...):** Cultural Exchanges
- **File 06 (Fact and Fiction):** Arts and Sciences

The subtitles below provide a detailed analysis of the treatment of each topic within each unit is provided. Thus, the selection of the materials, basically the listening and reading ones, is subjected to critique based on the authors' previously defined themes.

- **File 01: It's My Treat**

This file is characterized by its consistency in treating the theme '*Food and Drink*'. The first lesson "**Listen and Consider**" starts by introducing pictures of three dishes from different parts of the world (Chicken Tandori, India; Hamburger, USA; and Yorkshire pudding, England) and then focuses on the hamburger and the famous McDonalds restaurants. In the **Practice** section, other types of food and drinks are found (*write it up, p.20* in particular) such as *shorba, tea, artichokes, oranges, potatoes, sugar, coffee, mint...etc.* In "**Read and Consider**", the learners are supposed to discuss the notion of *recipe* first by discussing the ingredients needed to make *pancakes, tacharek, doughnuts, and Tamina* and then putting the instructions in the right order. A short text about the way to make pancakes is provided for the reading phase to prepare the learners for the coming activities in which they practice using the imperative to write instructions and both the imperative and sequencers to write a recipe. More food-related vocabulary is found in "**Words and Sounds**" for example *eggs, rice, peanuts, lamb, parsley, turnip, yoghurt, spaghetti, salmon, wheat, beans, peaches, lettuce...etc.* in addition to some cooking-related verbs, which are *to cook, to boil, to fry, to steam, to roast, and to grill*. Even the cartoon and the idioms in "**Take a Break**" are food-related. The cartoon is about *the fried fish meal* while all the idioms contain a word related to food or drinks like *she is as sweet as sugar*, which means she is very kind, and *it's not my cup of tea*, meaning I don't like it.

The consistency in treating the theme of *Food and Drink* in this unit is well illustrated in the last three lessons. "**Research and Report**" deals with three types of processed food (Cacao powder, chocolate bar, and cereal bar) and their ingredients and then presents Tony's diet on a typical day before asking the learners to write a letter to a British friend informing them about Algerian mealtimes, meals and what these meals are made of. A *menu* together with dialogue between a waiter and a customer at a restaurant are the materials selected by the authors to be the main content of "**Listening and Speaking**" to prepare the learners to practice

taking turns in similar situations. *Write it up* in this lesson requires the learners to write a note to describe the dish they will serve to an American friend visiting them. The same situation (at the restaurant) is present in "**Reading and Writing**" but with a focus on the shape and content of *a restaurant advertising leaflet*, which will be the topic of the project work that the learners are supposed to work on at the end of the unit.

- **File Two: You Can Do it!**

'*Citizenship and sustainable development*' is the theme that is intended to be treated in this file. Nevertheless, the contents of the file seem to be more or less selected based on the grammatical point of '*modal verbs*' that is reflected in the file's title (You can Do it), rather than being *topic-based*. Very few sentences (e.g. sentences c, d, f, g, i, and k. in act.4, p.50) and tasks (e.g act.5, p.53 and listen and check, p.54) in addition to the "**Project Round-up**" treat the theme set by the authors. The other lessons and parts of lessons deal with other topics. The materials and activities in "**Listen and Consider**" and "**Read and Consider**" are mainly about abilities, the school's context, and exam anxiety. However, the idioms and colloquialisms in "**Take a Break**" do not share the same topic but their selection seems to be made on the basis of expressing the functions of *ability/inability/possibility* using the modal verbs *can/cannot/may* respectively. The lesson of "**Research and Report**" is about the average life span of some animals and here too the focus is on the modal verbs *may/might/can't* to express degrees of possibility. Only the last group work activity of this lesson (act.5, p.53) is related to the topic of citizenship as it engages the learners in thinking as a member of the *World Wild Life Organization* and writing a ten-rule charter to ensure the protection of wild animals. The section of 'Listen and Check' in the next lesson "**Listening and Speaking**" discusses the same topic, i.e. protection of animals, whereas the second part of it focuses on another function, which is *agreeing* but the modal verb *can* is always present. "**Reading and Writing**", which is the last

lesson in the file, changes the topic suggesting a text within the domain of *artificial intelligence* talking about *robots* and the tasks and jobs they were/are/will be able to do.

- **File Three: Great Expectations**

The textbook authors have chosen the theme '*People and Places*' to be discussed in lessons of this file, and at the same time, many passages justify the choice of the file's title as they talk about the future plans and expectations of some people. Most of these passages are found in the listening scripts and reading texts in the lessons "**Listen and Consider**" and "**Read and Consider**" about the expectations of a brother and a sister, *Paul* and *Becky*. Within the same lessons, the activities and tasks in the *Practice* sections (p.69 & p.72) serve the same topic. Even the short texts in act.1 & act.6 in "**Words and Sounds**" talk about someone's expectations and future plans. Another concept, which is that of *superstitions* is introduced in "**Take a Break**". A set of expressions used in the British culture to expect what will happen in the future based on things that happen in the present are divided into two parts and put into columns to be matched by the learners and then they need to give their equivalents in the Algerian culture.

"**Research and Report**" deals with the second part of the theme '*places*' as it asks the learners to fill in the fact file about California and then report it to the class, in addition to making them draw the map of the U.S.A. and write each of the names of the states in full. Yet, the second part of the lesson deviates a lot from the file's theme as it discusses the educational systems in Algeria and the U.S.A. (a point that will be discussed nearly in the same way in unit 3 from *New Prospects SE3 textbook*). It also treats another topic found in *New Prospects (Unit 4)*, which is *advertising* in a task where it instructs the learners to fill in the fact sheet and then write *an advert* about the Royal Mauritanian Mausoleum but here it is closely related to the file's main theme '*people and places*'. "**Listening and Speaking**" is a typical lesson in terms of treating this latter by including a talk by a tourist guide addressing a group of tourists to

explain the program of a touristic visit to the Royal Mauritanian Mausoleum. "**Reading and Writing**", on the other hand, does not deal neither with the topic previously mentioned nor does it reflect the file's title. The reading material is in the form of the lyrics of a sad love song. It's only in the '*Write it out*' section that a link with the theme of '*people and places*' is found. Here, the learners are asked to write an email to a friend telling them about an excursion plan to a place of their choice and to anticipate solutions if some problems arise. The "**Project Round-up**" is about arranging a conference of boy or girl scouts or a local branch of *Young Friends of the Earth Society/Club*. It is clear that this project has no relation with the title of the file '*Great Expectations*', little to do with the theme '*people and places*', and better suits the theme of the previous file i.e. '*citizenship*'.

- **File Four: Then and Now**

The visual, listening and reading materials in this file do partially justify the choice of its title as almost all of them talk about people's past lives, ancient civilizations, and some famous historical, scientific and political figures in history. The comparison with the *Now* or the present is delayed to the end of the file i.e. in the "**Project Round-up**", where the learners are asked to make a poster about differences between past and present *customs and mores* (*food, clothes, ways of travelling, ways of celebrating weddings and religious festivals*) in their society. This is to be done by making interviews with their grandparents and observing how people live around them. Hence, the theme '*customs and mores*' chosen for this file does not seem to be discussed explicitly. The teachers are supposed to make efforts and strive to help their learners link the contents of the file with its theme.

To start with, the before-listening phase in "**Listen and Consider**" presents pictures of some actors and singers (e.g. Brad Pitt, Whoopi Goldberg, Arnold Schwarzenegger) and invites the learners to ask and answer questions about what they did for a living before they became stars. In the listening phase, an interview with Goldberg at a radio station about her previous

job as a high school teacher and her experiences with her students is to be discussed. Then, the learners are invited to talk about the things they used to do when they were younger through a questionnaire and a writing task. **"Read and Consider"** takes the learners back in time to the era of *the stone age* (life in the Sahara 5000 years ago) and how cavemen used to live, hunt, and manage to satisfy their daily life needs. However, **"Research and Report"** could be a typical lesson of the previous file dealing with *'people and places'* since the learners are supposed to fill in a fact file again but this time about *India*, carry out research to make a short history file about *India, Pakistan, and Bangladesh* and write a short presentation about *Taj Mahal*. In the same lesson, they are asked to make an internet search about the Abbasids Caliph *Harun al-Rachi* to write his biography and also about one of these people: *Al khawarizmi, Alhazen of Basra, Avicenna of Bokhara, Averroes of Cordova* in order to write an informative report about him.

"Listening and Speaking" and **"Reading and Writing"** shed the light on some parts of American history by dealing with the native Americans or *the Indian tribes*, who used to live in different parts of the American continent, and the political figure *Martin Luther King Junior* and his defence of civil rights in a peaceful way respectively. In the same respect, the *'Write it out'* section is about writing a short biography and a short letter of opinion about *Mahatma Gandhi* (India) and *Abdelhamid Ben Badis* (Algeria), who are two influential figures in their countries.

- **File Five: Dreams, Dreams...**

This file, according to the book's syllabus, is supposed to discuss the topic of *'cultural exchanges'* but the only lesson, in which it is explicitly treated, is **"Research and Report"**. This latter makes the learners collect data about *Australia* and some of the famous places, animals, and sports it is known with. **"Read and Consider"** and **"Listening and Speaking"** partially with the file's theme. The first lesson is about an Algerian (*Abdelkader*) and a British (*Ronald*)

talking about the things they are going to do soon, among which is the invitation of Abdelkader to Ronald to visit him in his parents' home in Tamanrasset. In "**Listening and Speaking**", Ronald has a conversation with a travel agent to plan for the trip to Tam but then he changes his destination to Ghardaia and from there makes a round trip to Ouargla, Biskra, Laghouat and then gets back to Ghardaia in three days before flying back to America. In the same lesson, the learners are invited to perform dialogues playing the roles of a taxi driver advising a tourist passenger and recommending the things to do in their town.

The learning materials as well as the activities and tasks in "**Listen and Consider**" and "**Reading and Writing**" reflect the file's title rather than the theme it is supposed to discuss. There, the learners will know about the *dreams* of *Bachir* and *Nacera* respectively and they will also practice using the second conditional to talk about *unreal situations (imaginary)* in the future.

The contents of the remaining lessons of the file i.e. "**Words and Sounds**", "**Take a Break**" and "**Project Round-up**" are related to some of the previously discussed points but they have nothing to do with the file's title and theme. The "**Project Round-up**", in particular, which is supposed to be the synthesis of the whole file, is neither about *dreams* nor about *cultural exchanges* as it assigns the learners to layout a newspaper problem page.

- **File Six: Fact and Fiction**

The topic chosen by the authors for the last file of the textbook is '*Arts and Sciences*' but as a matter of fact, the only lesson where this theme is present is "**Research and Report 2**" on page 152. In this lesson, learners are supposed to solve some brainteasers based on their knowledge of natural sciences in addition to discussing what the scientists *Archimedes*, *Isaac Newton*, *Joseph John Thomson*, and *Alexander Fleming* were doing when they made their discoveries. All the visual, reading, and listening materials together with the activities and tasks in all the other lessons seem to be selected based on their relevance to the file's title. To start

with, "**Listen and Consider**" is about a real situation of a daughter attacked by a shark while her family were on the beach. Similarly, "**Read and Consider**", is about another real situation (*fact*), which is that of an accident scene. The newspaper report implemented as reading material describes the scene and provides some details about the accident (e.g. time, place, actors, conditions...etc.). Other examples of *facts* are found in "**Words and Sounds**" and "**Take a Break**" on pages 149-150 respectively.

The second part of the file shifts the learners' attention to dealing with the second part of the title i.e. *fiction*. This starts with "**Research and Report1**" on page 151, where the first part of one of *Joha's* famous stories, *Joha and the Pots*, is given to the learners and then they are supposed to finish the story by asking old relatives about it or conducting internet research. In "**Listening and Speaking**", the learners are not really exposed to the notion of *fiction* since it is present only in the section of '*Your turn to speak*' by taking turns to act out a dialogue about the movie of the fictional character *Shrek*. However, "**Reading and Writing**" is entirely about *fiction* by making use of the famous and interesting story of '*Snow White*' as a reading material based on which the learners are supposed to do a set of activities and tasks. Finally, the "**Project Round-Up**" is a typical one for this file allowing the learners to make a scrapbook by choosing one among the four following topics:

- 1- A short report about a news item (e.g. an accident, a new invention)
- 2- An obituary of a celebrity who has just died.
- 3- A transcript in English of a short folktale told by Arabic- and/or Berber-speaking people.
- 4- A cartoon strip with speech bubbles illustrating the short folktale in task three above.

- **Adaptation of 'On the Move' to the Second-Generation Curriculum**

Some coursebook adaptation measures and procedures were conducted following the orienting guidelines of the *second-generation curriculum* for the use of 'On the Move', exceptionally, in the 2018-2019 school year instead of MBOE4, which has not been published yet. These measures, taken by the ministry of national education, are listed and explained in the teacher's guide: Middle School Key stage 3/Year Four (Ministry of National Education, 2018, pp.6-7). While the procedures are listed below, the textbook adaptation chart is joined to the appendices (**See appendix C**):

- a) An in-depth scrutiny of the whole coursebook was undertaken by a team of inspectors.
- b) Only the topics, texts, listening scripts, topical lexis, grammatical structures and functions as well as pronunciation activities and some communicative tasks **that fit the learning plan and help construct the competencies were selected. Note that 70% of the coursebook content has been kept; irrelevant topics were totally omitted.**
- c) To cope with the lack of relevant materials that fit the topics of some sequences, supplementary texts have been suggested to support teachers. **Needless to add that in terms of “fitness for purpose” teachers can adapt them or use others.**
- d) The in-depth analysis of “On the Move” revealed that most, if not all, of the activities, are mechanical; they are not tasks that should stimulate and push learners to exploit much of their intellectual and communicative abilities. **Therefore, the teachers can either create their own tasks or transform the coursebook’s activities into communicative tasks.**
- e) The selected content has been thematically categorized according to the suitable sequence.
- f) A guiding chart (**see Appendix I**) has been designed to help teachers find the right page of the coursebook content that should be included within the target sequence.

g) To facilitate its use, it has been decided that the chart should preferably contain the four sequences, **each of which should concomitantly be used with the coursebook's pages of the corresponding topics, texts, tasks and linguistic resources.**

4.1.2.1.2. New Prospects (SE3)

Each of the six units of the coursebook has a title next to the unit number preceding the learning/teaching content but also discusses a theme around which the unit's content is organized. These themes are found in the teacher's book as well as in the Learning Progressions English Year 3 (2018) and are listed as follows:

- 1- Unit One: Ancient civilizations
- 2- Unit Two: Ethics in business
- 3- Unit Three: Education in the world: comparing educational systems
- 4- Unit Four: Advertising, consumers and safety
- 5- Unit Five: Astronomy and the solar system
- 6- Unit Six: Feelings and emotions

• Unit One: Exploring the Past

The theme treated in this unit, as indicated above, is "Ancient Civilizations" and this is well reflected in the project outcome, which is about making *a profile* of an ancient civilization. The unit proposes a rich and varied content of civilizations from different continents and periods of history:

Ancient Greek, Roman, Ottoman and Western civilizations (Europe), Ancient Egyptian, Phoenician, Carthaginian and Numidia civilizations (Africa), Islamic civilization (Asia, Africa and Europe), Aztec, Maya and American civilizations (America), and Chinese, and Sumerian civilizations (Asia).

The first lesson "**Listen and Consider**" proposes a map and a radio interview about five (5) ancient civilizations, which are: Ancient Greek, Ancient Egyptian, Sumerian, Indus

Valley, and Chinese civilizations. It also contains two short texts about the Native Americans and their contribution to the American civilization in addition to the Maya civilization.

The main reading material in "**Read and Consider**" is a text entitled 'Algeria at the Crossroads of Civilizations', which is adapted from *Africana: The Encyclopedia of African and African American Experience*. The text gives a quick overview of the different civilizations and conquests that Algeria witnessed; from prehistoric civilizations in the Algerian Sahara, in particular, to the French colonization. The same lesson includes some examples talking about the Islamic civilization in Andalusia in addition to some short texts having as topics the Aztec and Maya civilizations, the Egyptian hieroglyphics and the influence of civilizations on one another through history.

"**Think, pair and share**" lesson asks the learners to write an *expository* essay about *the challenges faced by modern civilizations*; whereas "**Take a Break**" suggests a cartoon of a guy wearing sports clothes by different universal brands saying that he hates sport while another person commenting saying that he is one the many contradictions of our modern civilizations:



It also provides some proverbs and sayings related to civilizations, peace and war.

The Islamic civilization is present again in "**Research and Report**", where the learners are asked to research its achievements in different disciplines: medicine, astronomy, architecture...etc. by including pictures too.

"**Listening and Speaking**" includes a map illustrating the territories occupied by the Phoenician civilization in addition to a listening activity about the same theme. It also suggests a writing task asking the learners to prepare a short account of the development of Western civilization. However, the last lesson of the unit "**Reading and Writing**" refers back to the Ancient Egyptian civilization in an extended text as well as the Ancient Greek civilization in the writing task, in which students are required to follow some guidelines and to order the given illustrations (Trojan Horse, Ulysses Homecoming, Greek Siege of Troy, ...etc.) to write their version of *the myth of Ulysses*.

- **Unit Two: Ill-Gotten Gains Never Prosper**

'*Ethics in Business*' is the theme to be discussed in this unit. Many issues related to this theme are tackled and a huge set of vocabulary used to discuss such a topic is introduced. "**Listen and Consider**", for example, provides definitions for the terms *bribing, whistleblowing, lobbying, and false accounting*. Many other terms, such as *corruption/anti-corruption, tax evasion, speculators, anti-smuggling laws, public funds, investments, nepotism, bureaucratic abuse, fraud, money laundering, auditing, illegal payment, sleaze, embezzlement...etc.*, appear in the examples and dialogues included in the lesson. "**Read and Consider**" discusses *counterfeiting and fraud*, which are two important issues directly linked to the main theme of the unit '*Ethics in Business*'. The principal material of the lesson is a reading text entitled *Imitating property is a theft* adopted from *The Economist*, May 17th, 2003. The discussion of the topic of counterfeiting goes further in "**Think, pair, share**" as the learners are required to write an opinion article to disagree with people who argue that counterfeiting benefits consumers by giving them access to lower-price goods.

Unlike the contents of the lessons discussed above, the cartoon implemented in "**Take a break**" does not really match the theme of the unit. The cartoon shows a patient laying on a bed and a lot of medicines next to him. The people around him had the dialogue below:



The joke in the cartoon is a good one but there is no clear link between its content and the theme of the unit '*Ethics in Business*'. The proverbs and sayings, on the other hand, have *money* as a topic: *money talks*, *money is a good slave but a bad master*, *money doesn't grow on trees...etc.* however, the proverb *money can't buy you love* seems to be out of the unit's main concern. Only the last proverb '*business is business*' directly relates to it.

The contents of the last three lessons of the unit "**Research and Report**", "**Listening and Speaking**" and "**Reading and Writing**" are closely related to the topic '*Ethics in Business*' as they discuss: frauds and ethical issues in some practices, ethical standards of businessmen and their application in their factories and/or companies, and social auditing as both a social obligation and a necessity in doing business respectively. The unit's main project work is about writing a *charter of ethics* talking about *ethical standards* as an important dimension in all professions. Two alternative projects are suggested, which are:

- A report about internet piracy/cyber criminality: this one discussed ethical issues outside the domain of business.
- A report about counterfeits currently marketed in the local country.

- **Unit Three: Schools: Different and Alike**

The learning/teaching materials in this unit are selected according to their relevance to the unit's theme '*education in the world*'. The listening material in the lesson "**Listen and Consider**" is in the form of a conversation between a secondary school headmaster and a pupil's father. The other materials (examples, texts, dialogue) are around the *baccalaureate* examination, school and education in general. However, the pictures in the '*getting started*' section of the lesson illustrate different occupations to help the learners answer some discussion questions. This discussion together with the pictures may have an impact on learners' future careers as they target the post-baccalaureate phase. The five pictures represent the two genders in the following occupations: welder and cook for males, biologist, musician, and chemical engineer for females. Hence, males are basically coined with vocational training occupational roles while females are portrayed as university graduates from different disciplines. The reading text in "**Read and Consider**" has the title of *Education in Britain* but it refers in some of its paragraphs to the American educational system as to compare and contrast them. It also asks the students, in one of the activities (act.3, p.85), to compare primary or middle schools and the secondary school in Algeria based on some points (e.g. classes, students' age, subjects...etc.). At the end lesson, the learners are given guidelines to write a letter to a friend to describe their classroom/school and finally a cartoon representing a teacher with a pupil, who gives a funny answer identifying the zebra as a horse with a football shirt. "**Take a Break**" keeps the learners within the domain of education as it suggests a crossword game asking the learners to take the letters in italics and put them into the right order to spell out the school subjects.

The second part of the unit remains within the same theme '*education in the world*' starting with "**Research and Report**", which includes a picture illustrating a girls' secondary school in Britain in the 1960s and gives the students the task of designing a home page for their 'lycée' (secondary school). The first part of "**Listening and Speaking**", before listening,

shows a picture of *Marie Curie*, describing her as the *lady* who discovered *radium* but this is not true as the discovery was shared with her husband *Pierre* as she worked in his laboratory and they won both together the *Nobel Prize in Physics* in 1903. The remaining part of the lesson is about her life story/biography being a famous figure in the domain of science in general. The lesson moves suddenly, in the *after-listening* section, to ask the learners to write a lecture about *education in India* using some notes, and concludes with a writing task requiring the students to imagine their ideal school and then describe it with the help of a spider map. The textbook authors have chosen a text about the educational system in the U.S.A. as a reading material in "**Reading and Writing**" preceded by a diagram describing it and a question about the differences and similarities between the educational systems in Algeria and the U.S.A. The lesson concludes with a writing task about writing an *expository article* for a school magazine *analysing the causes and effects of examination stress on students*, which is a very important topic for the students at this phase as they prepare for the Baccalaureate examinations. The unit's project aims at designing an *educational prospectus* comparing and contrasting the British and Algerian educational systems by selecting a level (preferably secondary or higher education). Other alternative project outcomes that are closely related to the unit's theme are suggested:

- A leaflet for University Open Day,
- A portfolio of academic correspondence,
- A short educational guidebook for students.

• **Unit Four: Safety First**

Unit four discusses the topic of '*Advertising, consumers and safety*'. By the end of the unit, learners are supposed to work on a project in the form of a survey on the impact of advertising. They start by designing and administering a questionnaire to end up writing a report including different sections. The learning/teaching materials, examples, illustrations and

vocabulary included in the five main lessons of the unit are well-selected and are perfectly linked to the aforementioned theme. To start with, the interview between the *journalist* and the *market researcher* in "**Listen and Consider**" is about the impact of advertising on consumers' preference for organic food, being healthier, at the expense of conventional and genetically modified food (GMF). The same topic is triggered in the examples provided in the activities and tasks of the unit as well as the two short texts (act.1, p.110: benefits of garlic; and task p.111: consumers' safety) and the short dialogue (p.112), which is about healthy food. The textbook authors inserted a self-developed reading text in the lesson "**Read and Consider**" entitled '*How is your Energy Balance*'. It discusses mainly the contribution of advertising in changing people's eating habits. The lesson sticks to the theme of the unit '*advertising, consumers and safety*' in the different activities and tasks it includes. For example, the sentences in act.2, p.117 are under the title '*Health Warning*' and so is the writing task in '*Think, pair, and share*', which is about *food safety*. However, the text in act.2, p.118 and the dialogue in act.1, p.119 are about consumers, goods and prices. "**Take a Break**" and "**Research and Report**" contents are also about *food proverbs, food contents and health* (assignment one, p.122) and *home safety rules* (assignment two, p.122) respectively.

"**Listening and Speaking**" and "**Reading and Writing**", on the other hand, refer back to the topic of advertising. The two lessons discuss mainly the power of influence that advertisements have on consumers' choices by encouraging the learners to give and defend their opinions using arguments.

- **Unit Five: Are we alone?**

'*Astronomy and the Solar System*' is the theme around which the selection of the illustrations, listening and reading materials in this unit is based. A rich repertoire of technical vocabulary used in the field of *science and technology* is presented in the lessons' listening scripts, texts, tasks and activities. Examples of these technical terms are *telecommunication*

satellites, earth station, signals, telescope, planet, solar system, orbit, astronaut, space missions, spacecraft, gravity, astronomy, astrology, astrophysicist, astronomer, comet, galaxy, space exploration, space laboratories...etc. At the end of the unit, learners are invited to work on designing an astronomy booklet as their main project outcome. The textbook authors suggested other topics as alternative project works, which are closely related to the theme of the unit:

- A wall sheet about the solar system;
- A poster about the evolution of astronomy throughout the ages,
- A short sketchbook about major astronomers/astronauts,
- A short guidebook for amateur astronomers (How to use a telescope, how to observe the constellations ...).

The first part of "**Listen and Consider**", i.e. before and while listening phases, discusses the topic of '*telecommunication satellites*', which orbit the earth, send and receive signals from earth stations and at the same they can carry telescopes. The post-listening phase contains some activities, which stick to the same topic (e.g. act.3, p.140); whereas other activities and tasks contain sentences/short texts, which deal with planets of the solar system (e.g. act.2, p.138 & act.3, p.140) and so is the case of **Think, pair, share** in page 141. The reading material entitled '*The Solar System*' together with its pre-reading and post-reading activities in "**Read and Consider**" deal with the same topic too, i.e. planets of the solar system. To have a short break, as the title of the lesson reflects it, "**Take a Break**" portrays a boy with his Dad as earthmen on Mars in a cartoon having the following dialogue:



It also tries to make the learners use some idiomatic expressions containing the words *star* and *moon* such as *be born under an unlucky star* and *cry over the moon* in their appropriate context.

The second part of the unit does not deviate from its main theme as "**Research and Report**" assigns learners to work on either:

- An ID card of one of the eight historic celebrities (*Ptolemy, Caliph Al-Ma'mun, Al-Battani, Razez, Copernicus, Galileo, Hubble, Brahe*) to emphasize their respective *contributions to astronomy*;
- Research about *solar and lunar eclipses*.
- Finding the origin of some words related to astronomy (*nadir, zenith, and almanac*), adding four or five words belonging to the same origin and explaining the meaning of these words to their classmates;
- Or research about the myths about the heavens narrated by the ancestors or read about in other cultures.

The lesson of "**Listening and Speaking**" discusses *Extraterrestrial Intelligence*, as a debatable topic treated differently by *science* and *science-fiction*. However, "**Reading and Writing**" deals with a very interesting phenomenon in the field of astronomy, which is that of *comets*. The writing tasks in both lessons aim at training learners on how to *defend* an idea (saying it in writing p.155) and *refute* another (writing development p.160).

- **Unit Six: We are a Family!**

We are a Family is the title of a unit, which according to the teachers' guide, deals with the theme of *feelings and emotions*. The learners are exposed to a variety of materials and are asked to do a set of activities and tasks to conclude by working on a project that reflects the outcome of their learning. This latter is about writing a booklet of tips for coping with strong emotions. Learners can also opt for one of the following alternative projects:

- Writing a love poem or short story;
- Writing a short collection of jokes based on strong emotions;
- Making a sketchbook about major characters in love stories;
- Or adapting and acting out a scene or an act from a well-known play.

"Listen and Consider" in this unit does not focus on one single topic. Rather, it triggers a set of topics, which are independent of each other; yet, they are all linked to the unit's main theme. It discusses:

- the art of comedy, the sense of humour and the morals that comedies teach to people beyond the feeling of happiness they feel by watching them;
- the idea of the hero/heroine (idol);
- and seeking advice from *the agony aunt* to overcome the exam anxiety.

"Read and Consider" goes straight to the main idea of the unit by exposing the learners to a text entitled *Feelings*. It is adopted from *The Oxford Guide to British and American Culture* and compares the ways the British and the Americans express feelings and emotions. In addition to that, the lesson contains some activities discussing the same topic, for example, act.2, p.177 (friendship), act.2, p.178 (anger). Perhaps, the most interesting task regarding the topic of this research is the one in **"Think, pair, share" (p.180)**, which asks the learners to write a short newspaper article describing *the feelings and emotions expected from Algerian women and men in various situations*. Such an activity attracts the attention of learners from both sexes because

the learners tend to know about the other gender's reactions, feelings and emotions in different contexts and during the presentation phase, an atmosphere full of curiosity, debate, and controversial opinions risks to be dominating the class. "**Take a Break**" includes two sections, as usual, but in this unit, the authors started with the *idiomatic expressions*, which are to be matched with the feelings they correspond to. The second section invites the learners to transform the funny situation in the cartoon into a joke.

Moving to the second part of unit six (We are a Family!), nearly all the materials and activities in "**Research and Report**", "**Listening and Speaking**" and "**Reading and Writing**" discuss the same topic, i.e. *feelings and emotions*. There are only two remarks to be mentioned in this respect:

- *Saying it in writing* on page 187 is about preparing a speech to eradicate violence in sports and to call for fair play, which is not the topic of this unit.
- The reading material in the lesson "**Reading and Writing**" (entitled: *The Unicorn in the Garden*) is about a funny story of a husband, who lived happily ever after when the police and the psychiatrist took his wife and shut her up in an institution because of the story of the unicorn. This text is closely related to the title of the unit and also to the topic of the present research due to the debate it may create between learners from both genders about the way it pictures them, especially at the end of the story.

4.1.2.2. Activity/Task-Based Evaluation of “On the Move” and “New Prospects”

4.1.2.2.1. What is the learner expected to do?

The following tables summarise the analysis findings and present the percentages of activities and tasks within the two textbooks that correspond to each option. These percentages are calculated out of the total number of activities and tasks within each textbook.

A/ Take Turn

Initiate		Scripted response		Not required	
OTM	NP	OTM	NP	OTM	NP
15.3%	8.8%	73.7%	81.8%	11.0%	9.4%

Tasks that require a scripted response from the learners to direct questions using language supplied by the task represent the highest percentages in both textbooks with 73.7% in ‘On the Move’ and 81.8% in ‘New Prospects’. In the second place, tasks that require the learners to initiate using their own language with no prompt from the task represent 15.3% in ‘On the Move’, but only 8.8% in ‘New Prospects’, which makes them in the third place after tasks that require no response from the learners with a percentage of 9.4%. The learners in these tasks are not supposed to take any interactive role, and represent 11% of the total number of activities and tasks in ‘On the Move’.

B/ Focus 1: Language Components

Vocabulary		Grammar		Pronunciation		Morphology		Discourse	
OTM	NP	OTM	NP	OTM	NP	OTM	NP	OTM	NP
16.9%	15.9%	30.3%	31.8%	13.7%	9.4%	1.3%	7.7%	55.0%	34.7%

The analysis shows that discourse tasks have the highest percentage in both textbooks followed by those, which focus on grammar and language structures. The percentages of

activities, which focus on vocabulary and lexis, are in the third position followed by pronunciation and finally morphology activities. It is noticed that a very low percentage of these latter appears in ‘On the Move’ with only 1.3%.

Focus 2: Language Skills

Listening		Speaking		Reading		Writing	
OTM	NP	OTM	NP	OTM	NP	OTM	NP
11.3%	10.5%	18.3%	9.1%	19.6%	17.0%	14.7%	12.8%

The percentages of the tasks dealing with the four language skills are, more or less, balanced compared with the previously discussed ones. However, the reading tasks have the highest percentages in both textbooks. The catching remark is that the focus on speaking tasks is higher in ‘On the Move’ (18.3%) compared to ‘New Prospects’ (9.1%). The other remark is that higher percentages of tasks that focus on the other skills are present in ‘On the Move’ in comparison with ‘New Prospects’.

C/ Cognitive / Thinking Skills

Knowledge		Comprehension		Application		Analysis		Synthesis		Evaluation	
OTM	NP	OTM	NP	OTM	NP	OTM	NP	OTM	NP	OTM	NP
8.8%	10.5%	23.1%	22.2%	40.8%	24.1%	13.7%	20.5%	11.3%	12.5%	2.9%	10.8%

Tasks that require application skills are present with higher percentages in both textbooks compared to those, which require other cognitive skills. Similarly, lower-order thinking skills (i.e. knowledge, comprehension, and application) taken together have higher rates than higher-order thinking skills (i.e. analysis, synthesis, and evaluation). In ‘On the Move’, 72.7% of the activities target lower-order thinking skills compared to 27.9% target higher-order ones. In ‘New Prospects’, these represent 56.8% and 43.8% respectively. This

implies that tasks that trigger higher-order thinking skills are present in ‘New Prospects’ more than in ‘On the Move’. Percentages at the two ends of the hierarchy (i.e. knowledge and evaluation) are the lowest in both textbooks.

4.1.2.2.2. Who with?

Learners individually simultaneously		Pair work/ group work		Learner to whole class		Learner individually outside the class	
OTM	NP	OTM	NP	OTM	NP	OTM	NP
76.4%	73.3%	19.6%	21.9%	3.8%	10.2%	0.3%	0.0%

The majority of the tasks in both textbooks, according to the data in the table, require the learners to work individually and simultaneously inside the classroom whereas the tasks that make the learners work individually outside the classroom are totally absent in ‘New Prospects’ and present but with a very low percentage (0.3%) in ‘On the Move’. Pair work activities and tasks are found in both textbooks with close percentages (19.6% in OTM and 21.9% in NP). Yet, there is a noticeable difference between the percentages of the activities and tasks, in which the learner interacts with the whole class (3.8% in OTM and 10.2% in NP).

4.1.2.2.3. With what content?

A/ Input to learners

Extended discourse: Written		Extended discourse: Aural		Words/phrases/ sentences: Written		Words/phrases/ sentences: Aural		Graphic	
OTM	NP	OTM	NP	OTM	NP	OTM	NP	OTM	NP
27.1%	35.2%	9.9%	10.5%	53.6%	52.6%	2.9%	1.4%	13.4%	14.2%

The nature of the majority of input content provided to the learners within the activities and tasks is in the form of written words, phrases or sentences (53.6% in OTM and 52.6% in

NP) as opposed to aural ones, which represent the lowest percentages (2.9% and 1.4% respectively). The percentages for the extended discourse are not very different as written discourse input is present more than aural one in both textbooks. The graphic input is found in 13.4% of ‘On the Move’ and 14.2% of ‘New Prospect’s tasks and activities.

B/ Expected output from learners

Extended discourse: Written		Extended discourse: Oral		Words/phrases/ sentences: Written		Words/phrases/ sentences: Oral		Graphic		Other: Stress- intonation- matching...etc.	
OTM	NP	OTM	NP	OTM	NP	OTM	NP	OTM	NP	OTM	NP
14.2%	11.6%	10.2%	4.5%	32.2%	53.1%	21.2%	14.5%	1.1%	2.3%	12.3%	6.8%

The results of the analysis illustrated in the above table about the expected output from the learners are somehow different from those of the suggested input. In ‘New Prospects’, 53.1% of the activities and tasks expect output from the learners in the form of written words, phrases or sentences compared to 32.2% in ‘On the Move’. Activities and tasks that expect an oral output in the form of words, phrases and sentences come in the second position in both textbooks (21.2% in OTM and 14.5% in NP). The lowest percentages are noticed in tasks that expect a graphic output (1.1.% and 2.3% in OTM and NP respectively). More written extended discourse is expected from the learners than the oral one. Activities and tasks that expect a different output from the learners like identifying the stressed syllables, marking intonation, or matching between words, for example, are present with the following percentages: 12.3% in ‘On the Move’ and 6.8% in ‘New Prospects’.

Section Two: Thematic and Items' Analysis of English Language BEM and BAC Exam Papers

4.2.1. The BEM English Language Exam

In the BEM exam, the examinees are given one unique topic to be answered in no more than one hour and a half. The coefficient of the English language in the BEM exam is 2 from the total number of coefficients, which is 27. Hence, English is in the fifth rank together with natural sciences, physics and Islamic education and lagging behind Arabic (5), maths (4), French and History & Geography (3).

4.2.1.1. Description of the Exam Paper

English as foreign language exam papers, for both BEM 2014 and 2019, are identical in terms of time allocation (one hour and a half), format and number of activities. The only difference is in the title given to part two; Written Expression (BEM 2014) and Situation of Integration (BEM2019). They share the following structure:

- Text
- Part One:

A/ Reading Comprehension (3 activities / 07pts.)

- Two activities related to the comprehension of the text (05 pts.).
- One activity related to Lexis (02 pts.).

B/ Mastery of the Language (3 activities / 07pts.)

- One activity related to mechanics or morphology (02 pts.)
- One activity related to syntax (03 pts.)
- One activity related to pronunciation (02 pts.)

- Part Two: Written Expression/Situation of Integration (6pts.)

4.2.1.2. Analysis of the Exam Papers

4.2.1.2.1. Topics

BEM 2014 exam paper is mainly based on the content of “**Research and Report**” from File 03 of ‘On the Move’, which discusses the theme of *‘People and Places’*. The text, which is in the form of a report entitled “**BRAZIL**”, provides some information about the country. It also seems to justify the choice of this country in the first sentence, which states that Brazil is organising the Football World Cup taking place 12th June - 13th July 2014 (i.e. few days before the exam took place) and in particular because Algeria participated in it. This is enhanced in part two (Written Expression), as the examinees are asked to write a letter to a Brazilian friend, who seeks information about Algeria being one of the countries participating in the 2014 World Cup. They have to talk about Algeria’s situation, area, population, the capital city, main cities, and the things Algeria is famous for.

However, BEM 2019 English language exam paper deals with *food* and *eating habits* in different parts of the world. This seems to be inspired by File 01 of ‘On the Move’, which discusses the theme of *Food and Drink*. The suggested text has no title but it was adapted from “Food around the World” (Headway Oxford, 2004). Accordingly, Part Two (i.e. Situation of Integration) requires the exam takers to write an email to reply to an English pen-friend telling him/her about the eating habits in Algeria. The email needs to talk about the different kinds of Algerians eat, Algerian eating manners and some Algerian traditional dishes.

4.2.1.2.2. Exam Items’ Types

The overall typology of suggested activities for the BEM exam is joined to the list of appendices (**See Appendix D**). Table 5.1 presents the types of activities/tasks that both BEM 2014 and BEM 2019 included in their different parts and sections in addition to the instruction for each activity:

Table 4.1.*Activity / Task Types in BEM 2014 & BEM 2019 English Language Examination*

		BEM 2014		BEM 2019	
PARTS		Task types	Instructions	Task types	Instructions
PART ONE: READING	A/ Comprehension	Open-ended questions	Answer the following questions.	Open-ended questions	Answer the following questions.
		Information transfer	Complete the table with information from the text.	Information transfer	Complete the table with information from the text.
		Matching	Match words in column A with their equivalents from column B.	Matching	Match words in column A with their opposites from column B.
	B/ Mastery of the Language	Word order	Reorder the following words to get coherent sentences	Completion	Fill in the gaps with the suitable words.
		Transformation	Write the correct form of the verbs in brackets.	Transformation	Turn the following sentences into the negative.

		Sound recognition	Pick out from the text four words and complete the table according to the pronunciation of the final -s.	Sound recognition	Write the following words in the right box according to their final consonant sound.
PART TWO	Written Expression/ Situation of Integration	Writing a letter	Write him/her a letter telling him her about...	Writing an email	Write an email in which you tell him/her about Algerian eating habits.

The analysis of the table content reveals that the examinees dealt with almost the same types of activities. The only difference here is detected in *Mastery of the Language* (word order vs. completion). The other differences are at the level of instructions: Activities 2 and 3 from *Mastery of the Language* and *Part Two (Written Expression)*.

4.2.2. The Baccalaureate English Language Exam

Unlike the BEM exam, the examinees in the Baccalaureate are given two different versions for choice. Half an hour is added to the time allotted for the exam in order to give the examinees the chance to make the best choice based on the topic/theme and also the activities/tasks of each version in terms of type, difficulty and familiarity. The duration and coefficient of the English language examination in the Baccalaureate for both streams are presented in table 4.2.:

Table 4.2.

Duration and Coefficient of the English Subject in the BAC Exam

Streams	Duration	Coefficient
Foreign languages	03h, 30mins	05
Scientific	02h, 30mins	02

4.2.2.1. Description of the Exam Papers

The English language exam paper in the Baccalaureate for both streams; scientific and foreign languages, has the same general format with slight differences regarding the number of activities in Part One (Reading). These differences are found in cross-streams exam papers. However, some changes in the number of activities in Part One of the exam have been applied in BAC 2019 for both streams. The BAC examination paper revolves around the same theme and is organised as shown in tables 4.3. and 4.4.:

Table 4.3.*BAC 2014 English Language Exam Paper Organisation*

			Foreign Languages	Scientific
Part One:	Reading	Reading passage	150-200 words	120 – 150 words
	A. Comprehension	Number of activities	6 activities	5 activities
	B. Text Exploration	Number of activities	6 activities	5 activities
Part Two:	Written Expression	Number of topics (Choice)	2 (Guided/ free) 120-150 words	2 (Guided/ free) 80-120 words

Table 4.4.*BAC 2019 English Language Exam Paper Organisation*

			Foreign Languages	Scientific
Part One:	Reading	Reading passage	150-200 words	120 – 150 words
	A. Comprehension	Number of activities	5 activities	4 activities
	B. Text Exploration	Number of activities	5 activities	4 activities
Part Two:	Written Expression	Number of topics (Choice)	2 (Guided/ free) 80-120 words	2(Guided/ free) 80-120 words

4.2.2.2. Analysis of the Exam Paper

4.2.2.2.1. Topics

- **Scientific Stream**

'Ethics in Business', which is the theme of Unit Two, is present in 50% of the topics in BAC 2014 and BAC 2019 (i.e. 2 out of 4 topics). *Counterfeiting* and *cheating* being unethical behaviours are the topics of the second version of BAC 2014 and the first version of BAC 2019 respectively. However, the theme of Unit Four *'Advertising, Consumers, and Safety'* is discussed in the first version of BAC 2014, which has *consumer behaviour* as its main topic. The second version of BAC 2019 deals with the topic of *illegal immigration*, by including a text entitled 'Nightmare Journeys' that is directly linked to the theme of Unit Six *'Feelings and Emotions'*.

Whereas the first topics of Part Two: Written Expression, i.e. the guided ones, in all the versions follow their main topics, the second topics (free ones) deal with different themes. Three of these free topics ask the examinees to produce pieces of writing about *advertising and advertisements*, which is the theme of Unit Four (Safety First). These are found in the second version of BAC 2014 and both versions of BAC 2019. The free writing topic of the first version of BAC 2014 is about *space tourism* that has a relation with *'Astronomy and the Solar System'*, which is the theme of Unit Five (Are We Alone?).

- **Foreign Language Stream**

The theme of Unit Three *'Education in the World'* is present in 50% of the versions (i.e. two out of four). *Smoking at school* and *school violence* are the first topic of BAC 2014 and the second topic of BAC 2019 respectively. The second topic of BAC 2014 is about *life in Sparta*, which is related to the theme of Unit One *'Ancient Civilization'*; whereas the first version of BAC 2019 discussed the topic of *corruption* that is directly related to the theme of Unit Two *'Ethics in Business'*.

For Part Two of the exam papers (i.e. Written Expression), the guided task, as usual, discusses the same topic of the first part (Reading) but the free writing tasks have different topics allowing the examinees to make their choices based on their interest, amount of information, and level of mastery of the theme. The topics of the free writing tasks for the four versions of the Baccalaureate exam analysed here are displayed in Table 4.5.:

Table 4.5.

Free Writing Topics for BAC 2014 & 2019 (Foreign Languages Stream)

	VERSION	TOPIC	UNIT (THEME)
BAC 2014	One	Child labour	Unit Three (Education in the World)
	Two	Good behaviours when using social networks	Unit Two (Ethics in Business)
BAC 2019	One	Illegal immigration	Unit Six (Feelings and Emotions)
	Two	Counterfeiting	Unit Two (Ethics in Business)

4.2.2.2.2. Exam Items' Types

The typology of suggested activities that the Baccalaureate examinees, from both streams, are expected to face in the English language exam is found in the BAC Guide for the English Language Examination (October 2017) (**See Appendix E**). Nevertheless, the test items that were included in BAC 2014 and BAC 2019 are listed in Table 4.6. and 4.7.:

Table 4.6.

Activity / Task Types in BAC 2014 & BAC 2019 English Language Examination (Scientific Stream)

		BAC 2014 Scientific Stream		BAC2019 Scientific Stream	
PARTS		Topic 1 Tasks	Topic 2 Tasks	Topic 1 Tasks	Topic 2 Tasks
PART ONE: READING	A/ Comprehension	Type of discourse	Type of text	T/F statements	T/F statements
		T/F statements	T/F statements	Sentence order	Paragraph identification
		Comprehension questions	Comprehension questions	Comprehension questions	Comprehension questions
		Paragraph identification	Paragraph identification	Cohesive markers	Cohesive markers
		Cohesive markers	MCQ (title)	x	x
	B/ Text Exploration	(synonyms) Lexis	(opposites) Lexis	(opposites) Lexis	(synonyms) Lexis
		Chart completion (Morphology)	Word formation (Morphology)	Word formation (Morphology)	Word formation (Morphology)
		Asking questions about underlined words (Grammar)	Transformation (Grammar)	Combining sentences using connectors (Grammar)	Transformation (Grammar)

		Classifying words based on the number of syllables (Phonology)	Classifying words based on final – ed pronunciation (Phonology)	Reordering sentences (Discourse)	Dialogue completion (Discourse)
		Reordering sentences (Discourse)	Paragraph completion (Discourse)	x	x
PART TWO:	Written Expression	Guided 80-120 words composition	Free 80-120 words composition	Guided 80-120 words school magazine article	Free 80-120 words newspaper article
				Free 80-120 words composition	

By analysing Table 4.6., it can be noticed that the number of exercises was reduced in BAC 2019 to four (4) in both sections of Part One (Reading). Whereas there seems to be no specific activity type that was concerned by the reduction in the ‘comprehension’ section, the ‘phonology /pronunciation’ exercise was deleted from the ‘text exploration’ section in the 2019 Baccalaureate English language examination. Hence, the language-content exercises in BAC 2019 are limited to those of lexis, morphology, grammar and discourse. For the writing tasks, no changes were applied to the number of words the examinees are limited to in their pieces of writing, which is between 80 and 120 words regardless of the type/nature of the targeted discourse.

Table 4.7.

Activity / Task Types in BAC 2014 & BAC 2019 English Language Examination (Foreign Languages Stream)

		BAC 2014 (Scientific Stream)		BAC2019 (Scientific Stream)	
PARTS		Topic 1 Tasks	Topic 2 Tasks	Topic 1 Tasks	Topic 2 Tasks
PART ONE: READING	A/ Comprehension	T/F statements	Type of text	MCQ	T/F statements
		Sentence order	T/F statements	Paragraph identification	Paragraph identification
		Comprehension questions	Paragraph identification	Comprehension questions	Comprehension questions
		Type of text	Comprehension questions	Cohesive markers	Cohesive markers
		Cohesive markers	Type of discourse	Type of discourse	Title
		Title	Cohesive markers	X	x
	B/ Text Exploration	(synonyms) Lexis	(synonyms) Lexis	(opposites) Lexis	(synonyms) Lexis
		Word formation (Morphology)	Chart completion (Morphology)	Chart completion (Morphology)	Chart completion (Morphology)
		Transformation (Grammar)	Combining sentences using connectors (Grammar)	Combining sentences using connectors (Grammar)	Transformation (Grammar)
		Asking questions about underlined words (Grammar)	Tenses (Grammar)	Silent letters identification (Phonology)	Silent letters identification (Phonology)

		Classifying words based on the number of syllables (Phonology)	Classifying words based on final -s pronunciation (Phonology)	Reordering sentences (Discourse)	Reordering sentences (Discourse)
		Reordering sentences (Discourse)	Paragraph completion (Discourse)	X	x
PART TWO	Written Expression	Guided 120-150 words school magazine article	Free 120-150 words composition	Guide 80-100 words letter	Free 80-120 words web article
		Free 120-150 words composition	Guided 120-150 words composition		Free 80-120 words school magazine article

Table 4.7. shows that for the stream of the foreign languages, too, some changes were applied to the BAC 2019 English language exam paper in both parts. A ‘*comprehension*’ and a ‘*grammar*’ exercises were deleted. Therefore, the two sections of *Part One (Reading)* consist of five (5) exercises each instead of six (6) in BAC 2014. In *Part Two (Written Expression)*, the examinees are also required to limit themselves to a piece of writing that consists of 80-120 words instead of 120-150 words, which is the same number of words required from the scientific stream examinees.

4.2.3. Summary of the Findings

The findings of the BEM English language exam papers analysis show that these latter were very similar in terms of form and testing criteria and conditions (time allocation, scoring grid, number of exercises...etc.). Moreover, the topics dealt with were strictly linked to the themes of the two textbooks' units even if there were some textbook adaptations to the second-generation curriculum.

The English language exam papers in BAC 2014 and BAC 2019 were slightly different in terms of the number of exercises in Part One. One exercise is omitted from each of its two sub-parts (*Scientific stream*: 10 exercises in BAC 2014 / 08 in BAC 2019; *Foreign Languages Stream*: 12 exercises in BAC 2014 / 10 in BAC 2019). While in the part of 'Comprehension' there was no particular exercise type that has been triggered to be excluded from the grid, in BAC 2019, the '*Pronunciation/Phonology*' exercise and one of the two '*Grammar*' exercises were deleted from the scientific and foreign languages streams' exam papers respectively

The same textbook, '*New Prospects*' was used by the learners of both streams sitting for both exams. However, only Units Two and Six (i.e. Ill-Gotten Gains Never Prosper & We Are a Family) were dealt with by learners of both streams. The analysis of the topics of both exams for both streams has shown that the themes of the four units, that each stream is concerned with, were present but with differing rates of representation from one theme to another. '*Ethics in Business*' and '*Education in the World*' took the advantage in both exams for the scientific stream and the foreign languages stream respectively.

Conclusion

One major finding of the analysis conducted in this chapter is that the exam papers for both levels are strongly linked to the instructional materials used in the classroom in terms of the topics treated and the suggested activities. Learners examined in both exams are given the same chances of answering exam papers that follow a grid of exercises approved by the ministry of education and reviewed by a committee of experts before submitting them to the examinees. Although in the year 2019, learners were supposed to deal with the so-called '*second-generation curriculum*', the same textbooks were used in the final years of the middle and secondary school education, and the same format and content were found to exist in both exams administered in 2014 and 2019. These inferences are not exhaustive since further discussion will be present in *Chapter Six* of this thesis in order to link these findings with the results of the next chapter.

Chapter Five

Gender-based Analysis and Discussion of Learners' Academic Achievements in English Language BEM and BAC Exams

Table of Contents

Introduction	156
5.1. Research Methodology	156
5.1.1. Population and Sampling.....	156
5.1.2. Data Analysis Tools	157
5.1.2.1. z-Tests, <i>t</i> -Tests and the Central Limit Theorem	157
5.1.2.2. Cohen's <i>d</i> Effect Size	159
5.2. BEM English Language Examination	159
5.2.1. Population.....	159
5.2.2. Descriptive Statistics and Analysis	161
5.2.2.1. BEM 2014	161
5.2.2.2. BEM 2019	163
5.2.3. z-Tests/Effect Size Results and Interpretation.....	164
5.2.3.1. BEM 2014	165
5.2.3.2. BEM 2019	166
5.3. Baccalaureate English Language Examination.....	168
5.3.1. Population.....	168
5.3.2. Descriptive Statistics and Analysis	170
5.3.2.1. Scientific Stream	170
5.3.2.2. Foreign Languages Stream	173
5.3.3. z Tests/Effect Size Results and Interpretation	177
5.3.3.1. Scientific Stream	177
5.3.3.2. Foreign Languages Stream	180
5.3. Summary of the Findings.....	183
5.3.1. Descriptive Statistics	183
5.3.2. z-Tests.....	184
Conclusion.....	185

Introduction

This chapter aims to test the *null hypothesis* (H_0), which states that: *there is no difference between male and female learners' means in the BEM and BAC English language exams*. In this case, the research (alternative) hypothesis (H_1) is formulated as follows: There is a significant difference between male and female learners' means in the BEM and BAC English language exams. The chapter explains the data analysis procedures, provides the steps and assumptions for hypothesis testing, displays the tests' results and draws conclusions about the above-mentioned hypotheses.

5.1. Research Methodology

5.1.1. Population and Sampling

The population of this study consists of all learners sitting for the BEM and BAC exams for the years 2014 and 2019 in the Wilaya of Bordj Bou Arreridj. The Examinees' marks in the English language exams were collected from the official source (the local directorate of education). The choice of using the marks of these official exams is supported by the fact that:

- 1- They are national exams;
- 2- Learners relatively have the same test conditions;
- 3- The correction of the answer sheets is blind (anonymous papers)
- 4- Answer sheets are peer-corrected.

The changes in the curriculum (from the first to the second generation) and some other extraneous problems that the school years between 2014 and 2019 witnessed (teachers' strikes, exam topics leaked...etc.) were the reasons to opt for these exams' results.

For the BEM exam, all the learners followed a mainstream i.e. they use the same instructional materials and they are all examined based on a single and unique exam paper. Meanwhile, learners sitting for the BAC exam followed two different streams; the foreign

languages stream and the scientific stream. They have used the same textbook but dealt with different units, different learning objectives and focus, they had a different number of learning hours and different exam papers.

In order to conduct the z-tests (see 5.1.2.1.), a random sample, from each population chosen to take part in this research, was selected. Random sampling was done using the SPSS software based on 20% of each population. More details and also statistical data about the population sitting for each exam as well as the samples selected will be displayed in the coming sections.

5.1.2. Data Analysis Tools

To analyse the data, both Office Excel 2013 and SPSS (Statistical Package for Social Scientists) were used for the coding and the treatment of the learners' marks in the BEM and BAC English language exams for the years 2014 and 2019. In SPSS, the coding of the data was done on a numerical basis. Hence, male and female learners' were attributed the values '1' and '2' respectively. Data analysis was done using three tools:

- English Mark / Gender Cross Tabulation;
- z-tests;
- Cohen's *d* calculation of effect size.

The first method of grouping data has already been referred to in Chapter Three. Each of the other two procedures is explained in the following sections.

5.1.2.1. z-Tests, t-Tests and the Central Limit Theorem

Comparing two populations' means is very common. However, the use of a *t* statistic or a *z* statistic depends on whether the two populations' standard deviations are known or not. *Z* statistics assume that the population standard deviation (σ) is known and this is especially when the sample size is small (<30). They are also used when the sample size is large (>30) whether the populations' standard deviations are known or not. The value of (σ) is required to calculate

the standard error. The t statistic is used when the population standard deviation is unknown (notably in the case of a small sample size i.e. <30), which is most frequently the case and is estimated using the sample standard deviation. The formulas for the t and z statistics are similar and the only apparent difference is in the denominator, which is in z the standard error based on the population standard deviation ($\sigma/\sqrt{N}$) while for t , the denominator is based on the sample standard deviation ($s/\sqrt{N}$).

One very important theorem that is closely related to these tests, is *the central limit theorem*. According to OpenStax (2015), “The Central Limit Theorem only holds if the sample size is “large enough” which has been shown to be only 30 observations or more.” (p.286). Moreover, Nachmias & Guerrero (2018) further explain that:

The significance of the central limit theorem is that it tells us that with a sufficient sample size the sampling distribution of the mean will be normal regardless of the shape of the population distribution. Therefore, even when the population distribution is skewed, we can still assume that the sampling distribution of the mean is normal, given random samples of large enough size. Furthermore, the central limit theorem also assures us that (a) as the sample size gets larger, the mean of the sampling distribution becomes equal to the population mean and (b) as the sample size gets larger, the standard error of the mean (the standard deviation of the sampling distribution of the mean) decreases in size. The standard error of the mean tells how much variability in the sample estimates there is from sample to sample. The smaller the standard error of the mean, the closer (on average) the sample means will be to the population mean. Thus, the larger the sample, the more closely the sample statistic clusters around the population parameter. (pp. 299-300)

In our case, and since the two populations' standard deviations are known, z -tests are used, and the sampling distribution for the difference between the means is normal in accordance with the central limit theorem.

5.1.2.2. Cohen's d Effect Size

Cohen's d is a measure of effect size based on the differences between two means. It is used for the measurement of the magnitude of the difference between the two groups (Daniels & Minot, 2020). Using Cohen's d , the effect size is determined by calculating the mean difference between the two groups and then dividing the result by the pooled standard deviation:

$$\text{Cohen's } d = (\mu_2 - \mu_1) / \text{SD pooled}$$

Where:

$$\text{SD}_{\text{pooled}} = \sqrt{\left(\frac{SD_1^2 + SD_2^2}{2} \right)}$$

According to Cohen (1988) cited in Daniels & Minot (2020, p.152), effect sizes are defined as small when $d = 0.2$, medium when $d = 0.5$, and large if $d = 0.8$.

In this research, the calculation of the effect size was done with the help of a specialized social sciences statistics website: <https://www.statskingdom.com/effect-size-calculator.html>.

5.2. BEM English Language Examination

5.2.1. Population

The population of this part of the study are fourth-year middle school pupils in the Wilaya of Bordj Bou Arreridj, who sat for the BEM exam in the years 2014 and 2019. The choice of this population was not arbitrary since the textbook used as supporting material for the teaching/learning of English in the period from 2014-2019 was "On the Move". The difference is that in the school year 2018/2019, as has already been discussed in Chapter Four, 'On the Move' was adapted to fit the second-generation curriculum. Hence, the point was to discuss and compare the results of learners, who studied English based on two different

curricula. The following tables provide some descriptive data concerning the targeted population.

Table 5.1.

BEM 2014 Population Frequency

		Gender			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Male	6178	51,7	51,7	51,7
	Female	5765	48,3	48,3	100,0
	Total	11943	100,0	100,0	

Table 5.1. clearly shows that nearly 52% of the population, who sat for the English language exam in BEM 2014 (N=11943) are males ($N_1 = 6178$). The number of female learners ($N_2 = 5765$) at this level is about 4 % lower than that of males ($N_2 - N_1 = 409$).

Table 5.2.

BEM 2019 Population Frequency

		Gender			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Male	5699	47,8	47,8	47,8
	Female	6229	52,2	52,2	100,0
	Total	11928	100,0	100,0	

Table 5.2. displays descriptive data that is different from the previous one. The number of female learners ($N_2=6229$) is higher than that of males ($N_1=5699$) and the difference between the two sub-groups is, again, about 4% of the whole population ($N=11928$). The two populations' sizes are closely identified with a difference of 15 examinees only.

5.2.2. Descriptive Statistics and Analysis

5.2.2.1.BEM 2014

Table 5.3 displayed some statistics about the population of BEM 2014 examinees. The number of examinees classified based on their gender, the two categories means, standard deviations, medians, minimum and maximum scores as well as the ranges of their scores are presented.

Table5.3.

BEM 2014 Population Descriptive Statistics

Gender	Mean	N	Std. Deviation	Median	Minimum	Maximum	Range	Kurtosis	Skewness
Male	7.65	6178	4.20	7.00	.50	20.00	19.50	-.44	.58
Female	8.71	5765	4.50	7.50	.50	20.00	19.50	-.67	.54
Total	8.16	11943	4.37	7.00	.50	20.00	19.50	-.52	.58

According to Table 5.3., the males' population ($N_1=6178$) mean is $\mu_1= 7.65$ ($SD=4.20$); whereas, the females' population ($N_2=5765$) average of scores is $\mu_2=8.71$ ($SD= 4.50$). This seems to give a slight advantage to females. This applies too to the median, which equals 7.00 and 7.50 for males and females respectively. The other descriptive statistics (minimum score, maximum score, and range) are found to be identical.

Figures 5.1. is a histogram illustrating the grouped frequency distribution of male and female examinees' scores in the BEM 2014 English language examination.

Figure 5.1.

*English Mark * Gender Cross tabulation Chart (BEM 2014)*

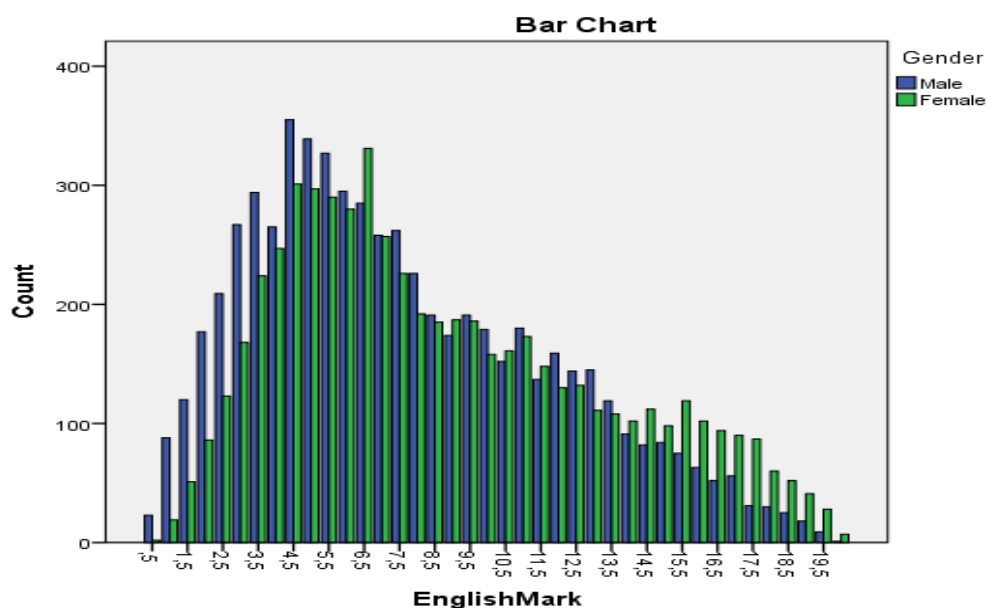


Figure 5.1. shows that the distribution of the marks is positively skewed for both categories. This is the case when the median is lower than the mean (see table 5.3). Since the skewness values are both between .50 and 1.00 ($Skew_1=.58$ / $Skew_2=.54$), both distributions are described as moderately skewed. Males seem to outnumber females at the left end of the chart (i.e. Marks ≤ 6); whereas, the number of female examinees is greater than that of males for all the marks that are equal or superior to fourteen (14) (Mark ≥ 14). Concerning the scores, which range between six (6) and fourteen (14), there is a kind of balance with a slight superiority of one gender over the other from one mark to another.

5.2.2.2.BEM 2019

Table 5.4.

BEM 2019 Population Descriptive Statistics

Gender	Mean	N	Std. Deviation	Median	Minimum	Maximum	Range	Kurtosis	Skewness
Male	7.29	5699	3.82	6.50	.50	19.50	19.00	.02	.71
Female	8.83	6229	4.39	8.00	.50	19.50	19.00	-.68	.44
Total	8.09	11928	4.20	7.50	.50	19.50	19.00	-.39	.59

Table 5.4. reveals that the males' population ($N_1=5699$) mean $\mu_1= 7.29$ ($SD=3.82$) is lower than the females' population ($N_2=6229$) average of scores $\mu_2=8.83$ ($SD= 4.39$). Again, this difference is in favour of females. The median of female scores ($MDN_2= 8.00$) is higher than males' one ($MDN_1=6.50$). The other descriptive statistics (minimum score, maximum score, and range) are shown to be identical as was the case with BEM 2014 population.

Figure 5.2.

*English Mark * Gender Cross tabulation Chart (BEM 2019)*

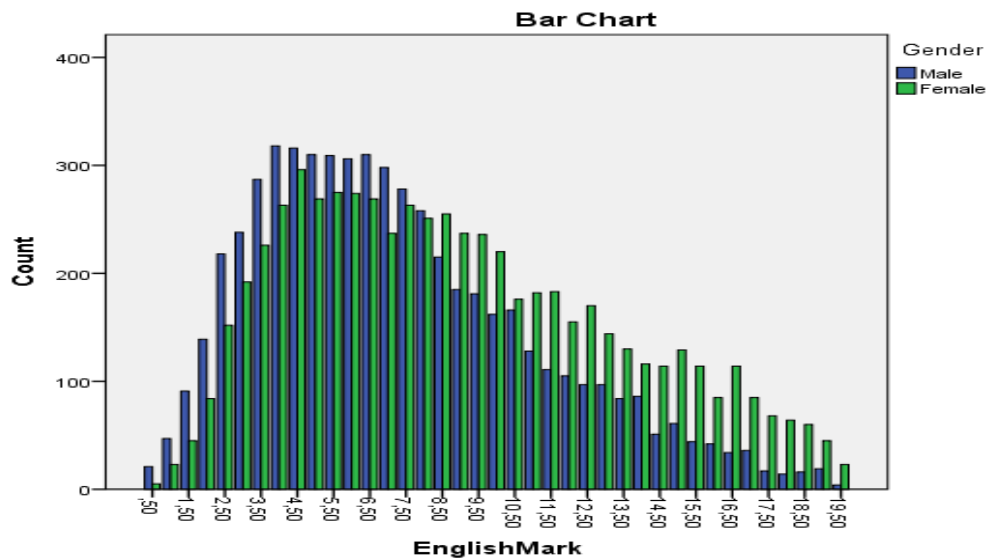


Figure 5.2, shows that the distribution of male and female scores is positively skewed just like the one in Figure 5.1. This is supported by the data in Table 5.4, where the skewness of the two groups' scores' distributions has positive values ($Skew_1=.71$ i.e. moderate skewness and $Skew_2=0.44$ i.e. approximately symmetric skewness). This is due, again, to the fact that the medians are lower than the means for both categories of the populations ($MDN_1=6.50 < \mu_1=7.29$ / $MDN_2=8.00 < \mu_2=8.83$). The chart can be analysed by splitting it into two parts, where the reference mark is seven and a half (7.50), which is the median of the whole population ($MDN=7.5$). There, the number of examinees from both genders is quite balanced. However, for all the other scores below eight (7.50) i.e. $mark < 7.50$, the number of males is greater than that of females; while the reverse situation is present for the scores above seven and a half (7.50) i.e. $mark > 7.50$.

5.2.3. z-Tests/Effect Size Results and Interpretation

Although the displayed figures provide a general description of the distribution of marks based on students' numbers from each gender, they cannot give statistically significant data based on which conclusions about gender differences in exam performance can be generated. Hence, the need to make use of two sample z-tests to compare male and female examinees'

scores in the BEM 2014 and BEM 2019 English language examinations was necessary to check the existence of a significant difference; and if that is the case, the magnitude of the difference between the means is calculated using Cohen's d coefficient of effect size.

The results of the two sample z -tests are displayed in tables including the samples' means, known variance, samples' sizes, hypothetical mean difference (set at 0, which refers to the null hypothesis that $\mu_1 - \mu_2 = 0$), the z -score, in addition to the p -value and the critical z -score. All these help in the interpretation of the result that is done immediately after each table.

The interpretation of the result consists of four parts: **1-** H_0 hypothesis testing, **2-** p -value (Type I error possibility), **3-** z -score statistic and **4-** Cohen's d effect size is calculated and reported to show the degree of the magnitude of the difference between the two populations' averages.

5.2.3.1.BEM 2014

- **Result**

Table 5.5.

z-test: Two Sample for Means BEM 2014

	<i>Variable 1</i>	<i>Variable 2</i>
Mean	7.506921824	8.767105263
Known Variance	16.9964883	20.9222269
Observations	1228	1140
Hypothesized Mean Difference	0	
z	7.023423672	
$P(Z \leq z)$ one-tail	1.08247E-12	
z Critical one-tail	1.644853627	
$P(Z \leq z)$ two-tail	2.16493E-12	
z Critical two-tail	1.959963985	

Interpretation

1. H_0 hypothesis

Since $p\text{-value} < \alpha$, **H_0 is rejected**.

The mean of **male examinees** is considered to be **not equal to** the mean of **females**.

In other words, the difference between the mean of the **male** and **female** populations is big enough to be statistically significant.

2. $p\text{-value}$

$p\text{-value}$ equals **2.16493E-12**. This means that the chance of a type I error (rejecting a correct H_0) is very small. The smaller the $p\text{-value}$ the more it supports H_1 .

3. The statistics

The test statistic z equals **7.023423672**, is not in the 95% critical value accepted range: [-1.9600: 1.9600].

4. Effect size

The effect size for this analysis ($d = 0.07$) was found to be below the range of Cohen's (1988) convention for a small effect ($d = 0.2$).

5.2.3.2.BEM 2019

Likewise, table 5.6., presents the results of the $z\text{-test}$ for male and female examinees samples' scores in the BEM 2019 English language exam.

Table 5.6.

$z\text{-test}$: Two Sample for Means BEM 2019

	Variable 1	Variable 2
Mean	7.210241	8.958195
Known Variance	15.04764	19.07717

Observations	1201	1208
Hypothesized Mean Difference	0	
Z	-10.3865	
P(Z<=z) one-tail	0	
z Critical one-tail	1.644854	
P(Z<=z) two-tail	0	
z Critical two-tail	1.959964	

- **Interpretation**

1. **H₀ hypothesis**

Since $p\text{-value} < \alpha$, **H₀ is rejected**.

The mean of **male examinees** is considered to be **not equal to** the mean of **females**.

In other words, the difference between the mean of the **male** and **female** populations is big enough to be statistically significant.

2. **p-value**

$p\text{-value}$ equals **0**. This means that the chance of type I error (rejecting a correct H₀) is very small.

3. **The statistics**

The test statistic z equals **-10.3865**, and is not in the 95% critical value accepted range: [-1.9600: 1.9600].

4. **Effect size**

The effect size for this analysis ($d = 0.1$) was found to be below the range of Cohen's (1988) convention for a small effect ($d = 0.2$).

5.3.Baccalaureate English Language Examination

5.3.1. Population

Third-year secondary school learners in the Wilaya of Bordj Bou Arreridj, who sat for the BAC exam in the years 2014 and 2019, represent the population of this part of the study. The researcher opted for the achievements of learners studying within two different streams, which are the scientific and foreign languages. This is done on purpose as the themes and topics dealt with in these latter are quite different and even the shared ones are not treated in the same way and depth. The aim is to deeply discuss the achievements of learners having different interests and learning backgrounds in addition to the comparative method, based on the gender variable that is adopted in carrying out the analysis of data in this part of the research. Both populations (i.e. BAC 2014 & BAC 2019) used 'New Prospects' as supporting material for learning English as a foreign language. Descriptive data concerning the population chosen for this part of the research are displayed in the following tables.

Table 5.7.

Frequency Table of BAC 2014 Population

Scientific Stream	Frequency	Percent	Valid Percent	Cumulative Percent
Male	2032	41,2	41,2	41,2
Female	2895	58,8	58,8	100,0
Total	4927	100,0	100,0	
Foreign Languages stream	Frequency	Percent	Valid Percent	Cumulative Percent
Male	167	28,9	28,9	28,9
Female	410	71,1	71,1	100,0
Total	577	100,0	100,0	

Table 5.8.

Frequency Table of BAC 2019 Population

Scientific Stream	Frequency	Percent	Valid Percent	Cumulative Percent
Male	1651	39,2	39,2	39,2
Female	2559	60,8	60,8	100,0
Total	4210	100,0	100,0	
Foreign Languages stream	Frequency	Percent	Valid Percent	Cumulative Percent
Male	304	31,8	31,8	31,8
Female	651	68,2	68,2	100,0
Total	955	100,0	100,0	

5.3.2. Descriptive Statistics and Analysis

5.3.2.1. Scientific Stream

Table 5.9.

BAC 2014 (Scientific Stream) Population Descriptive Statistics

Gender	Mean	N	Std. Deviation	Median	Minimum	Maximum	Range	Kurtosis	Skewness
Male	9.22	2032	3.30	9.00	.50	19.00	18.50	-.37	.35
Female	9.75	2895	3.35	9.50	1.00	19.00	18.00	-.51	.19
Total	9.53	4927	3.34	9.50	.50	19.00	18.50	-.47	.26

Table 5.9., shows that the males' population ($N_1=2032$) mean $\mu_1= 9.22$ ($SD=3.30$) is slightly lower than that of the female population ($N_2=2895$) $\mu_2=9.75$ ($SD= 3.35$). The median of males' scores ($MDN_1= 9.00$) is lower than females' one ($MDN_2=9.50$), which means that half of the males' scores (50%) are below 9.00, while for females, 50 % of the scores are below the score of 9.50. This does not seem to show a big difference. Concerning the other statistics, there is a difference in the range of scores ($R_1=18.50 / R_2=18.00$) since the minimum score differs (.5 for males and 1.00 for females).

Table 5.10.*BAC 2019 (Scientific Stream) Population Descriptive Statistics*

Gender	Mean	N	Std. Deviation	Median	Minimum	Maximum	Range	Kurtosis	Skewness
Male	8.85	1651	3.55	8.50	.00	19.50	19.50	-.33	.60
Female	10.27	2559	3.64	10.00	.00	19.50	19.50	-.72	.32
Total	9.71	4210	3.67	9.00	.00	19.50	19.50	-.62	.42

Table 5.10., indicates that the males' population ($N_1=1651$) mean $\mu_1=8.85$ ($SD=3.55$) is lower than the females' population ($N_2=2559$) average of scores $\mu_2=10.27$ ($SD=3.64$). This statistic, once more, conveys that females outperform males. The median of males' scores ($MDN_1=8.50$) is lower than females' one ($MDN_2=10.00$), which means that half of the males' scores (50%) are below 8.50 whereas for females the same rate is below the score of 10. For the other descriptive statistics: minimum score (.00), maximum score (19.50), and range (19.5), it is noticed that they are the same for both groups.

The comparison between male and female examinees' scores in BAC 2014 and BAC 2019 English language examinations based on the distribution of the examinees' scores is illustrated in the cross-tabulation charts in figures 5.3 and 5.4:

Figure 5.3.

*English Mark * Gender Cross tabulation Chart (BAC 2014, scientific stream)*

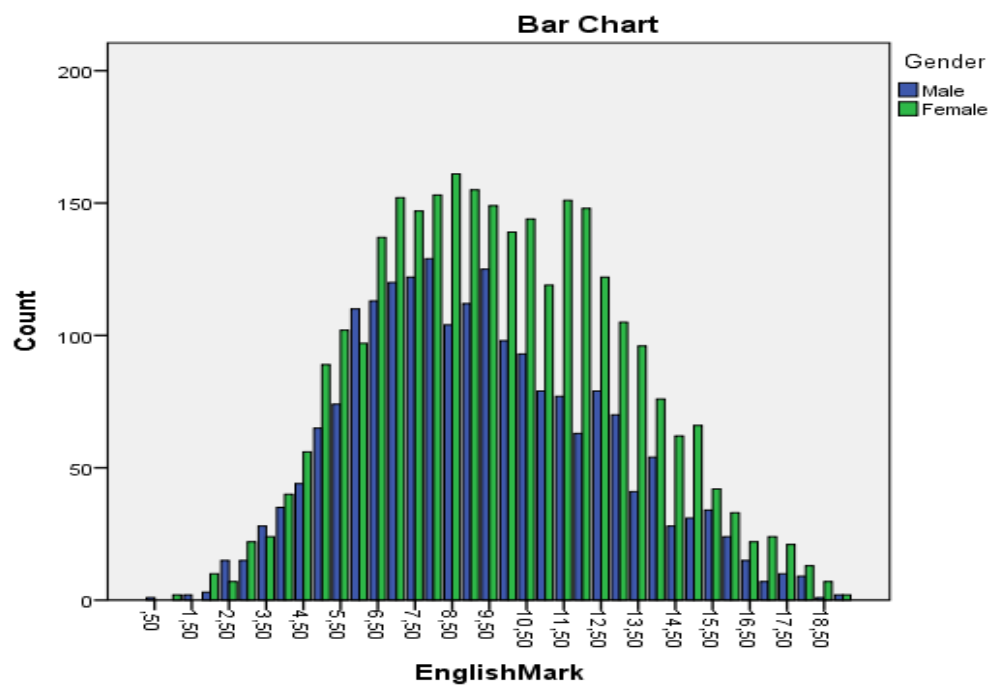
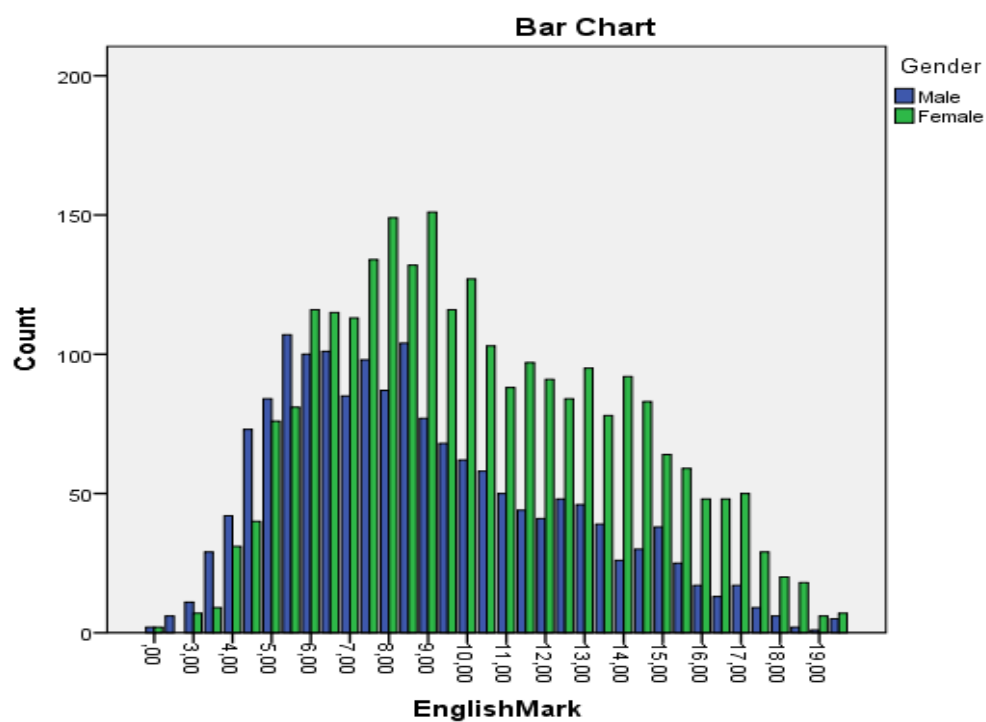


Figure 5.4.

*English Mark * Gender Cross tabulation Chart (BAC 2019, scientific stream)*



Both figures present positively skewed distributions with a long tail on the right. This is supported by the skewness values in Tables 5.8 and 5.9 respectively. Skewness for BAC 2014 scores' distributions ($Skew_1=.35$ / $Skew_2=.19$) is approximately symmetric (i.e. between -.5 and .5). On the other hand, the distribution in Figure 5.4 (BAC 2019) is approximately symmetric for female scores ($Skew_2=.32$); however, it tends to be moderate for male scores' distribution ($Skew_1=.60$ i.e. between 0.5 and 1). The figures display quite similar results, especially for the marks which are between 6.5 and 20, i.e. $[6.5, 20[$, where the number of female students for each score inside this range is greater than that of males in both exams. Concerning the scores that range between 0 and 6.5, i.e. $[0, 6.5[$, the results are not consistent in figure 5.3 (BAC 2014). Yet, apart from the score of six (6), males outnumber females in each of the scores within that range in the results displayed in figure 5.4. (BAC 2019).

5.3.2.2.Foreign Languages Stream

Table 5.11.

BAC 2014 (Foreign Languages Stream) Population Descriptive Statistics

Gender	Mean	N	Std. Deviation	Median	Minimum	Maximum	Range	Kurtosis	Skewness
Male	10.02	167	2.90	10.00	2.00	17.00	15.00	-.05	-.23
Female	10.33	410	2.74	10.00	3.50	19.00	15.50	-.12	.16
Total	10.24	577	2.79	10.00	2.00	19.00	17.00	-.06	.03

The male population's ($N_1=167$) mean $\mu_1= 10.02$ ($SD=2.90$) is nearly the same as that of the female population ($N_2=410$) $\mu_2=10.33$ ($SD= 2.74$). Surprisingly, the median of males' scores equals that of females' one ($MDN_1=MDN_2=10.00$), which means that half of males' and females' scores (50%) are below or equals the score of 10.00 and the same rate is above or equals that score. Moreover, there is a difference in the range of scores ($R_1=15.00 / R_2=15.50$) since the minimum scores ($Min_1=2.00 / Min_2=3.50$) and the maximum scores ($Max_1=17.00 / Max_2=19.00$) differ.

Table5.12.

BAC 2019 (Foreign Languages Stream) Population Descriptive Statistics

Gender	Mean	N	Std. Deviation	Median	Minimum	Maximum	Range	Kurtosis	Skewness
Male	9.17	304	3.10	9.50	1.00	18.00	17.00	-.54	-.03
Female	9.42	651	3.16	9.00	1.00	17.50	16.50	-.43	.30
Total	9.34	955	3.14	9.00	1.00	18.00	17.00	-.44	.20

Once more, the male population's ($N_1=304$) mean $\mu_1= 9.17$ ($SD=3.10$) is nearly the same as that of the female population ($N_2=651$) $\mu_2=9.42$ ($SD= 3.16$). The median of males' scores ($MDN_1= 9.50$) is, for the first time, higher than that of females' scores ($MDN_2=9.00$). Moreover, there is a difference in the range of scores and this statistic is also in favour of males ($R_1=17.00 / R_2=16.50$) since the minimum scores are equal ($Min_1=Min_2=1.00$) while the maximum scores ($Max_1=18.00 / Max_2=17.50$) differ.

The frequency distribution of male and female examinees' scores they obtained in BAC 2014 & 2019 English language examinations are illustrated in figures 5.5 and 5.6.

Figure 5.5.

*English Mark * Gender Cross tabulation Chart (BAC 2014, foreign languages stream)*

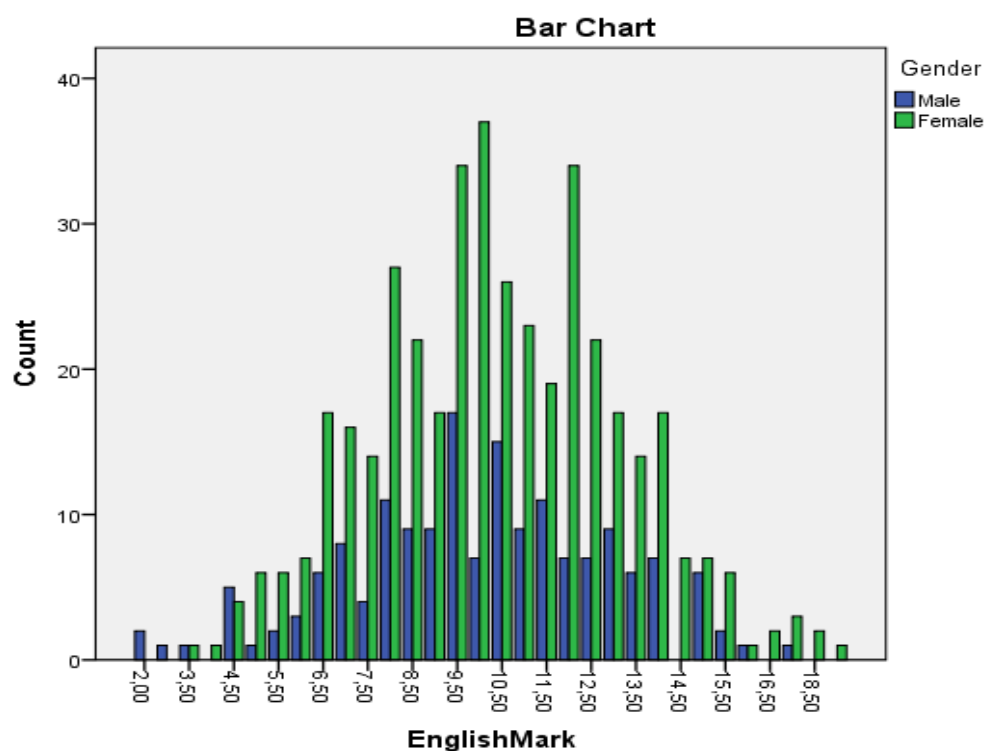
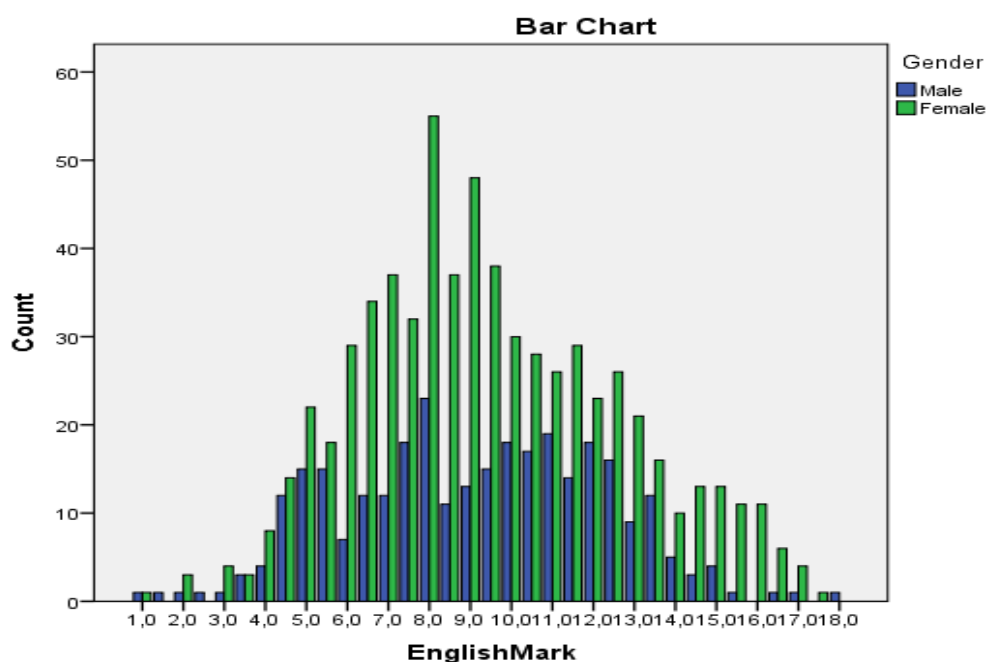


Figure 5.6.

*English Mark * Gender Cross tabulation Chart (BAC 2019, foreign languages stream)*



One noticeable remark regarding the male scores' distributions, compared to the previous distributions, is that they are negatively skewed. This is confirmed by the values ($Skew_1 = -.23 / -.03$) displayed in Tables 5.10 and 5.11 respectively, which means that the distribution is left-tailed. However, since the values are between $-.50$ and $.50$, the distributions are considered to be approximately symmetric. Female scores distributions are also approximately symmetric; yet, they are positively skewed i.e. the distribution is right-tailed ($Skew_2 = .16 / .30$).

In BAC 2014, the scores for the whole population of examinees ranged between two (2) and nineteen (19); [02, 19], however, they were between one (1) and eighteen (18); [01, 18] in BAC 2019. The first impression one may have when analysing the graphs is that females did better than their male counterparts as they outnumber them in almost every score within the two ranges; with some few exceptions, basically the scores lower than four and a half (4.5).

5.3.3. z Tests/Effect Size Results and Interpretation

5.3.3.1. Scientific Stream

In order to compare male and female examinees' score means in the English language examination in BAC 2014 & 2019 for the scientific stream, z-Tests for the selected random samples were used. The effect size is to be calculated in case there is a significant difference between the means. Table 5.13., presents the results of the two sample z-test for BAC 2014 male and female examinees scores:

Table 5.13.

z-test: Two Sample for Means BAC 2014 (Scientific Stream)

	<i>Variable 1</i>	<i>Variable 2</i>
Mean	8.964115	9.722222
Known Variance	10.73732	10.80654
Observations	418	567
Hypothesized Mean Difference	0	
z	-3.58386	
P(Z<=z) one-tail	0.000169	
z Critical one-tail	1.644854	
P(Z<=z) two-tail	0.000339	
z Critical two-tail	1.959964	

- **Interpretation**

1. **H₀ hypothesis**

Since $p\text{-value} < \alpha$, **H₀ is rejected**.

The mean of **male examinees** is considered to be **not equal to** the mean of **females**.

Alternatively, the difference between the mean of the **male** and **female** populations is big enough to be statistically significant.

2. **p-value**

$p\text{-value}$ equals **0.000339**. This means that the chance of type I error (rejecting a correct H₀) is very small.

3. **The statistics**

The test statistic z equals **-3.58386**, and is not in the 95% critical value accepted range: [-1.9600: 1.9600].

4. **Effect size**

The effect size for this analysis ($d = 0.07$) was found to be below the range of Cohen's (1988) convention for a small effect ($d = 0.2$).

Similarly, Table 5.14., displays the results of the two sample z -Test for BAC 2019 (Scientific Stream) scores:

Table 5.14.*z-test: Two Sample for Means BAC 2019 (Scientific Stream)*

	Variable 1	Variable 2
Mean	8.584286	10.49014
Known Variance	12.13972	13.51176
Observations	350	507
Hypothesized Mean Difference	0	
z	-7.69545	
P(Z<=z) one-tail	7.11E-15	
z Critical one-tail	1.644854	
P(Z<=z) two-tail	1.42E-14	
z Critical two-tail	1.959964	

- **Interpretation**

- 1. H_0 hypothesis**

Since $p\text{-value} < \alpha$, **H_0 is rejected.**

The mean of **male examinees** is considered to be **not equal to** the mean of **females**.

In other words, the difference between the mean of the **male** and **female** populations is big enough to be statistically significant.

- 2. p -value**

p -value equals **1.42E-14**. This means that the chance of type I error (rejecting a correct H_0) is very small.

- 3. The statistics**

The test statistic z equals **-7.69545**, and is not in the 95% critical value accepted range: [-1.9600: 1.9600].

4. Effect size

The effect size for this analysis ($d = 0.04$) was found to be below the range of Cohen's (1988) convention for a small effect ($d = 0.2$).

5.3.3.2.Foreign Languages Stream

Table 5.15.

z-test: Two Sample for Means BAC 2014 (Foreign Languages Stream)

	Variable 1	Variable 2
Mean	10.22727	10.55747
Known Variance	11.06392	7.310612
Observations	33	87
Hypothesized Mean Difference	0	
Z	-0.50993	
P(Z<=z) one-tail	0.305049	
z Critical one-tail	1.644854	
P(Z<=z) two-tail	0.610099	
z Critical two-tail	1.959964	

- **Interpretation**

1. **H₀ hypothesis**

Since $p\text{-value} > \alpha$, **H₀ cannot be rejected.**

The mean of **male examinees** is considered to be **equal to** the mean of **females**.

In other words, the difference between the mean of the **male** and **female** populations is not big enough to be statistically significant.

2. *p*-value

p-value equals **0.610099**. This means that if we would reject **H₀**, the chance of type I error (rejecting a correct **H₀**) would be too high: **0.610099 (61%)**. The larger the *p*-value the more it supports **H₀**.

3. The statistics

The test statistic *z* equals **-0.50993**, which is in the 95% critical value accepted range: [-1.9600: 1.9600].

4. Effect size

The observed standardized effect size is *very small* (***d* = 0.04**). That indicates that the magnitude of the difference between the means is small.

Table 5.16.

z-test: Two Sample for Means BAC 2019 (Foreign Languages Stream)

	Variable 1	Variable 2
Mean	8.905405	9.268595
Known Variance	10.42244	11.38767
Observations	74	121
Hypothesized Mean Difference	0	
Z	-0.74927	
P(Z<=z) one-tail	0.226847	
z Critical one-tail	1.644854	
P(Z<=z) two-tail	0.453693	
z Critical two-tail	1.959964	

- **Interpretation**

1. **H₀ hypothesis**

Since $p\text{-value} > \alpha$, **H₀ cannot be rejected.**

The mean of **male examinees** is considered to be **equal to** the mean of **females**.

In other words, the difference between the mean of the **male** and **female** populations is not big enough to be statistically significant.

2. **p-value**

$p\text{-value}$ equals **0.453693**. This means that if we would reject **H₀**, the chance of type I error (rejecting a correct **H₀**) would be high: **0.453693 (45.36%)**. The larger the $p\text{-value}$ the more it supports **H₀**.

3. **The statistics**

The test statistic z equals **-0.74927**, which is in the 95% critical value accepted range: [-1.9600: 1.9600].

4. **Effect Size**

The observed standardized effect size is *very small* (**$d = 0.03$**). That indicates that the magnitude of the difference between the means is small.

5.3. Summary of the Findings

5.3.1. Descriptive Statistics

To start with, the analysis of male and female examinees' scores in the English language examinations in BEM 2014 & BEM 2019 has shown that the results were quite similar (**BEM 2014:** $\mu_1 = 7.65$, $\mu_2 = 8.71$; **BEM2019:** $\mu_1 = 7.29$, $\mu_2 = 8.83$). The mean of the female examinees' scores is higher than that of the male one in both exams. The same observation could be noticed while comparing the medians in both exams' results (**BEM 2014:** $MDN_1 = 7.00$, $MDN_2 = 7.50$; **BEM2019:** $MDN_1 = 6.50$, $MDN_2 = 8.00$). Nevertheless, male and female examinees' minimum and maximum scores, and the range of scores, were identical. The distribution of the examinees' scores in BEM 2014 is positively and moderately skewed for both genders. This is the case also for the distributions of the scores in BEM 2019. The only difference is that the skewness of the distribution of the female examinees' scores is approximately symmetric.

Moving to the BAC exam, the mean of the scientific stream male examinees' scores was higher in BAC 2014 than that of BAC 2019 ($\mu_1 = 9.22$ / 8.85 respectively) while that of females was higher in BAC 2019 ($\mu_2 = 9.75$ / 10.27). The female means are higher than those of the males in both exams. However, the means of foreign languages stream examinees' scores were very close to each other (**BAC 2014:** $\mu_1 = 10.02$, $\mu_2 = 10.33$; **BAC 2019:** $\mu_1 = 9.17$, $\mu_2 = 9.42$). As can be noticed, there is always a small superiority of females over males in both exams.

This superiority can also be noticed in the medians of scores in both exams for the scientific stream (**BAC 2014:** $MDN_1 = 9.00$, $MDN_2 = 9.50$; **BAC 2019:** $MDN_1 = 8.50$, $MDN_2 = 10.00$). On the contrary, the results of the foreign languages stream have shown that the medians of scores for the two categories are the same in BAC 2014 ($MDN = 10.00$) and a reversed superiority was found in terms of the medians of scores in BAC 2019 ($MDN_2 > MDN_1$). Hence, half the number of male examinees' scores were above $MDN_1 = 9.50$, whereas the same number of females' scores was above $MDN_2 = 9.00$.

The distribution of the scores in BAC 2014 (Scientific Stream) is positively skewed for both categories in both exams, which means that the average is greater than the median. Yet, the degree of skewness differs from one category to another from one exam to another exam too. The skewness is approximately symmetric for the distributions of scores in BAC 2014 and that of females' scores in BAC 2019, but the distribution of males' scores in BAC 2019 tends to be moderately skewed.

For the foreign languages stream, one striking remark is that the distributions of male examinees' scores in both exams are negatively skewed, which means that the median is greater than the mean. Although the distributions are approximately symmetric, this observation is unique among all the analyses that have been conducted. Female examinees' scores distributions are also approximately symmetric in both exams but they are positively skewed.

5.3.2. z-Tests

The analysis of the two sample z-tests' results for the scores of the BEM exams revealed that the means of male and female examinees' scores were significantly different. Hence, the null hypothesis was rejected and the alternative hypothesis was confirmed. However, the effect size that is represented by Cohen's d equals 0.07 and 0.1 in BEM 2014 and BEM 2019 respectively, which means that the effect size can be considered to be trivial (i.e. below the range of Cohen's (1988) convention for a small effect).

The same result was found after the analysis of the scientific stream examinees' results in BAC 2014 and BAC 2019. They revealed that although male and female scores' means were significantly different in both exams, the effect size was very small ($d=0.07$ & $d=0.04$ respectively), which means again that the gender of the examinees did not greatly affect their performance on the English language exams.

Once more, the analysis of the results of the foreign languages stream examinees in BAC 2014 and BAC 2019 has given different conclusions from the previously discussed results.

As the p -value was greater than α ($\alpha=0.05$), the null hypothesis *could not be rejected*, which means the mean of male examinees is considered to be *equal* to that of female ones for the two population samples in both exams. In this case, the difference between the mean of the male and female populations is not big enough to be statistically different.

Conclusion

The analysis of male and female learners' achievements in all the exams shows that the effect of gender on scores is either not statistically significant (the case of foreign languages stream BAC exam results) or, even in the case where the difference was statistically significant, the size of the effect was trivial (BEM exams and scientific stream BAC exams). Moreover, foreign languages stream male examinees have shown relatively better performance in the English exam than that of females as their distributions were negatively skewed. In a trial to explain these conclusions, it can be noted that in the first case (i.e. foreign languages stream BAC exams), the English language subject is a fundamental one and its coefficient is five (5). This can justify the fact that there was no difference between male and female score means due to the importance of the subject, its exam and especially its exam outcome compared to the status of the same subject and its exam in the second context (the BEM exam and the BAC exam for the scientific stream). In this latter, English has two (2) as a coefficient and it does not have the status of a fundamental subject. Taken as a whole, then, we can conclude that females do not do better than males because they have better language capacities but because of some other extra factors apart from gender.

Chapter Six

General Discussion, Implications, and Recommendations

Table of Contents

Introduction	185
6.1. Exam papers vs. Textbooks: Content-based Comparison.....	185
6.1.1. BEM Exam Papers and ‘On the Move’	185
6.1.2. BAC Exam Papers and ‘New Prospects’	188
6.2. Discussion of the Statistical Data.....	192
6.2.1. The BEM Exam Statistics	192
6.2.2. The BAC Exam Statistics.....	193
6.3. Summary of the Discussions	195
6.4. Recommendations and Implications	198
6.4.1. To the Policy Makers and Curriculum Developers	198
6.4.2. To the Syllabi and Textbook Designers	200
6.4.3. To the Exams’ Designers	201
6.5. Limitations of the Study.....	202
6.6. Suggestions for Further Research	202
Conclusion.....	203

Introduction

In this chapter, a general summary and a cross chapters' discussion of the main findings of the textbooks' and exam papers' analysis in addition to the statistical data results are provided. In other words, it is the place, where the three research questions are thoroughly answered and also, where the three research hypotheses are being analysed and validated/refuted at once. This is going to be done step by step in order to arrive at the sum of the conclusions at last. Moreover, this chapter suggests some implications both in the pedagogical and educational domains, provides some recommendations to some essential actors in the educational field and ends up by giving some suggestions for further research taking into account the contribution the present study offers to the fields of education, research and gender studies.

6.1.Exam papers vs. Textbooks: Content-based Comparison

6.1.1. BEM Exam Papers and 'On the Move'

To start with, and prior to dealing with the content-based discussion, it is likely to mention that there was an implied change in the exam design philosophy at the level of Part Two, which changed from *written expression* in BEM 2014 to *situation of integration* in BEM 2019. This change of terminology is, basically, due to the second-generation curriculum reforms and renewal.

The themes discussed in both BEM exam papers (i.e. 2014 & 2019) were inspired by 'On the Move' textbook content. BEM 2014 is all around the topic of *People and Places*, which, according to the fourth-year middle school syllabus, fits within the content of File 03 from the textbook. Similarly, the topic discussed in BEM 2019 was *food around the world* which is coherent with File 01's theme: *Food and Drink*.

In terms of activity types, the BEM exams were almost identical and they strictly stuck to the BEM Guide for the English Language Examination guidelines. In fact, the only difference between the two exam papers was found in Part One: Reading, B/ Mastery of the Language, Activity One, where in BEM 2014 the examinees were asked to reorder words to get coherent sentences (**word order task**), while in BEM 2019 they were supposed to fill in the gaps with the suitable words (**completion task**).

Apart from *Part Two: Written Expression/Situation of Integration*, where the examinees were supposed to use their language to write a paragraph, all the other activities in both exam papers require from them a scripted response to some provided prompts. This matches with the textbook's analysis concerning the distribution of activities in which the learner has to initiate (15.3%), respond to questions using language required by the task (73.7%) or where no response is required at all using the language (11%) in the textbook but no activity in the exams.

The analysis of the activity types based on the targeted language components revealed that the majority of activities and tasks deal with discourse either written or spoken (55%), followed by grammar (30.3%), vocabulary (16.9%), pronunciation (13.7%) and finally morphology (1.3%). Worth mentioning here is that one single activity may focus on both, for example, discourse and grammar or vocabulary and pronunciation at the same time. Moving to the analysis of the activities on the basis of language skills they focus on, the results have shown that reading tasks are in the first position with 19.6% of the total number of activities followed by those which focus on speaking (18.3%), then writing (14.7%) and finally listening (11.3%). This implies that the activities that have a focus on discourse are oral more than written-oriented. Comparing this with the classification of the exam activities, the first catching remark is that no speaking section is included in the test. Second, no morphology exercise is suggested. Third, discourse exercises are represented by *Part Two*

in addition to activity one (Part One: B/ Mastery of the Language) in both BEM exams - obviously done in a written form- with a total number of eight (8) points. Grammar and pronunciation are present with only one exercise of each but with an advantage of three (3) points for the grammar exercise and two (2) points for the pronunciation one. Concerning *Part One*, devoted to reading comprehension, the first exercise is in the form of short-answer items, the second is an information transfer one (completing a table from the text), and the third exercise is about vocabulary (matching words with their synonyms/opposites), which is out of two (2) points in both exams. Therefore, apart from the absence of speaking and a morphology exercise, the distribution of the points in the BEM exam seems to be logical compared to the representation of exercises with the different focuses discussed before in the textbook.

The analysis based on the cognitive/thinking skills targeted by the textbook activities revealed that lower-order thinking skills are more targeted than higher-order ones (72.7% and 27.9% respectively). Something, which needs reconsideration since, at least, a balance between the levels needs to be achieved in order to enhance the learners' higher-order cognitive skills. Yet, the situation in the exam papers does not seem to be different as the only task, in which the examinees are supposed to make some synthesis is in *Part Two: Written Expression/Situation of Integration*. The first section of *Part One: Reading* involves learners in practising their *comprehension* skills in activities one and two while the third activity is at the lowest level - *knowledge* – as it requires them just to match between words that are synonymous or opposite in meaning. The second section suggests two *application* activities, which are activity two (action verb: turn /write) and activity three (action verb: write/pick up), whereas activity one also from the learners to *analyse* the input and either order the words or fill in the gaps (*analysis* level). This makes five (5) activities fall in the

lower-order level and only two (2) in the higher-order level and once more this reflects the textbook's content.

In both BEM exams, the input content that is provided to learners starts with an extended written discourse (text) followed by activities that also provide written input but in the form of words, phrases or sentences. Hence, there exists no aural or graphic input in both exam papers. Compared to the textbook content, where 80.7% of the activities and tasks provide such input, this seems reasonable since the only two targeted skills in the exam are *reading* and *writing*.

Likewise, apart from activity three from *Mastery of the Language (Part One: Reading)*, in which the examinees are asked to match between words, the expected output from the examinees in both BEM exams takes a written form. This can be either as an extended discourse (*Part Two: Written Expression/Situation of Integration*), or words, phrases or sentences as is the case in the remaining activities. Again, since the exams do not target listening and speaking skills, there is no expected oral output from the examinees and no graphic one as well. When it comes to the 'On the Move' textbook, the results have shown that the percentage of the activities and tasks, in which the learners are expected to produce a written output (46.4%) is relatively not very different from those, which expect an oral one (31.4%) in both forms (i.e. discourse/words, phrases, sentences).

6.1.2. BAC Exam Papers and 'New Prospects'

The examinees sitting for the BAC exam are given the chance to work on one topic out of two. Hence, the total number of exam papers that were analysed was eight (8): two BAC exams (2014&2019), two streams, and each stream with two exam versions for the examinees to choose one of them. The topics treated in all the exam papers without exception were relevant to the textbook content. Worth mentioning that two unit themes were present

in 50% of the exam topics: *Ethics in Business* (Unit Two) and *Education in the World* (Unit Three).

The analysis of the English language BAC exam papers has revealed that the examinees are required to give scripted responses to direct questions using language supplied by the activity/task throughout the whole exam with the only exception of *Part Two: Writing*, in which they can use their language in the free topic mainly. This is reflected to a great extent in ‘New Prospects’, where such kind of activities/tasks (scripted response) is present with a percentage of 81.8% i.e. giving the learners the freedom to initiate using their language in only 8.8% of the total number of the textbook’s exercises and tasks.

In terms of language components, it has previously been shown in Chapter Four that *discourse* and *grammar* activities are more present in the textbook (34.7% and 31.8% respectively) than the activities focusing on the other language aspects (*vocabulary*: 15.9%; *pronunciation*: 9.4%; and *morphology*: 7.7%). For the scientific stream, all the language components are present in the exam paper in BAC 2014. However, in BAC 2019 the number of exercises was reduced in section two of *Part One: Reading* to four (4) instead of five (5) and the language component that was eliminated from the exam was *phonology*. This could be joined to the absence of any speaking or listening sections in the exam to reach conclusions and implications about the English language exam policy and ideology as well as the resulting outcomes. It is almost the same situation in the foreign languages stream exam papers. The only difference is that the number of exercises in the *Text Exploration* section (Part One: Reading) was reduced from six (6) to five (5) and instead of two (2) grammar exercises in BAC 2014, only one (1) grammar exercise was suggested in both versions in BAC 2019. Hence, all the language components were present in all the exam versions, with a focus on grammar in BAC 2014.

Related to what has been discussed in the previous paragraph, the focus of the English language BAC exam in terms of language skills is only on *reading* and *writing* just like in the BEM exam discussed earlier. This is reflected in ‘New Prospects’ since the activities and tasks dealing with the *reading* skill have the highest percentage (17%) followed by those related to *writing* (12.8%) while *listening* and *speaking* are present in only 10.5% and 9.1% of the total number of activities and tasks respectively. This reflects an unbalanced distribution of activities and tasks targeting the four skills in the textbook in addition to totally discarding the assessment of learners’ speaking and listening abilities, which are essential to determine someone’s mastery of the language.

The statistics of the textbook analysis have shown that the activities and tasks targeted in ‘New Prospects’ target lower-order thinking skills (approximately 56%) more than higher-order ones (approximately 44%). As in ‘On the Move’, the lowest percentages are at the bottom and top of the hierarchy i.e. *knowledge* (10.5%) and *evaluation* (10.8%) and to a lower extent *synthesis* (12.5%). The majority of the activities and tasks target the cognitive skills of comprehension, application and analysis with a total percentage of approximately 67%. In the English language BAC exam, a whole section (*Part One: Reading, Section A*) is devoted to *comprehension* and suggests 4-6 exercises (BAC2014/2019 and Scientific/foreign languages streams) that all target that cognitive skill. *Part Two: Writing*, in all the BAC exam papers, is also aimed at targeting one single cognitive skill, which is *synthesis*. Here, the examinees are asked to write a composition about one particular topic. The exercises in section B of *Part One: Reading*, which is *Text Exploration*, vary between those targeting *knowledge* (with action verbs like *find*), *application* (verbs like *complete* and *combine*), and *analysis* that seems to be more represented with activities asking the examinees, for example, to *classify*, *re-order*, and *divide*.

The majority (=73.3%) of the activities and tasks in the third-year secondary school textbook 'New Prospects' require the learners to work individually and simultaneously inside the classroom. Although this limits the amount of their interaction with their classmates through pair, group, or whole class work, this seems to train them to take their official exam at the end of the year, as they would face similar conditions.

The input content provided for the learners in the BAC exam papers is very similar to that in the BEM ones. The exam begins with an extended written discourse (a text) followed by a set of activities and tasks that provide the learners with a similar type of input but in the form of words, phrases or sentences that the examinees would use to give their answers. Again, similar to the BEM exam, there exists no aural or graphic input content in all BAC exam versions. Compared to the input content of the activities and tasks in 'New Prospect', where 87.8% of the input is in a written form, it can be concluded that more focus is put on written language rather than aural reception or graphic representations in either material i.e. instructional and testing ones.

All that has been discussed above apply to the activities and tasks' expected output in both exam papers and the instructional textbook 'New Prospects'. The examinees are supposed to give written answers to all the exam activities with no expected oral or graphic responses. Similarly, the expected output from the learners while dealing with the textbook's tasks is mainly in a written form with an approximate percentage of 65%. Oral output is expected from 19% of the textbook's tasks, which is relatively a very low percentage for a language learner to practise their speaking and only 2.3% of the output is expected to be in a graphic form. Other types of output e.g. stress and intonation identification, matching etc. are expected from 6.8% of the activities.

6.2. Discussion of the Statistical Data

6.2.1. The BEM Exam Statistics

The descriptive statistics for both BEM 2014 and BEM 2019 have shown a slight advantage of females over males in terms of the average score (mean) and the median of scores. As a result, this was reflected in the distributions of marks that were illustrated in Figures 5.1. and 5.2. (See Chapter Five), which show that the curve is positively skewed for both categories and in both exams. This occurs whenever the median is lower than the mean and reveals that the majority of the marks for male and female examinees are depicted below the mean of scores. In BEM 2014, both score distributions were moderately skewed as the skewness values were between .50 and 1.00 ($Skew_1=.58$ / $Skew_2=.54$), which is the same case for the male scores' distribution in BEM 2019 ($Skew_1=.71$), but for females, the score distribution's skewness was approximately symmetrical i.e. between -.50 and .50 ($Skew_2=0.44$). The standard deviation values are lower for the male population in both exams (BEM 2014: $SD_1= 4.20$ / $SD_2= 4.50$; BEM 2019: $SD_1= 3.82$ / $SD_2= 4.39$). This implies that the values are spread out over a wider range in the females' scores compared to males' ones.

On the other hand, the inferential statistics conducted using data related to the BEM exam, aimed, basically, at testing the first part of the second null hypothesis of this research, which was set as follows: *There is no difference between male and female learners' means in the BEM and BAC English language exams.* The focus here is obviously on learners' means of scores in the BEM English language exam. The two sample z-tests results comparing male and female examinees' means in both BEM 2014 and BEM 2019 have shown that the p -value is lower than α ($p\text{-value} < \alpha$). Therefore, the null hypothesis was rejected and the research (alternative) hypothesis stating that: *There is a significant difference between male and female learners' means in the BEM and BAC English language*

exams was confirmed in its part related to the BEM English language exam. In BEM 2014, the p -value equals **2.16493E-12** (i.e. divided by 1000000000000, which makes it a very small value); whereas, in BEM 2019, the p -value equals 0. These results reflect that the risk of committing a type I error i.e. rejecting an H_0 (null hypothesis) while it is true is very small.

The fact that the null hypothesis is rejected does not provide much information about the magnitude of the difference between the two groups. Although the results of the two samples' z -tests have revealed a significant difference between male and female means, the calculation of Cohen's d has shown that the difference is negligible even if it is significant since both d -values are less than 0.2 (BEM 2014: $d = 0.07$, BEM 2019: $d = 0.1$). Hence, we can conclude that: *even if the null hypothesis was rejected, the magnitude of the difference between the two groups' means is very small and there seems to be no significant gap based solely on the gender factor.*

6.2.2. The BAC Exam Statistics

To start with, the descriptive statistics for the scientific stream examinees' scores in both exams (i.e. 2014 & 2019) have revealed that females had higher mean scores. This was the case for the median of scores too. And since all the medians were lower than the means for both categories and in both exams, the curves, illustrated in figures 5.3 and 5.4 (**See Chapter Five**), are both positively skewed with differing degrees of symmetry. This is supported by the values of skewness in Tables 5.9 and 5.10. The standard deviation values are quite similar for the two populations in both exams (BAC 2014: $SD_1 = 3.30$ / $SD_2 = 3.35$; BAC 2019: $SD_1 = 3.55$ / $SD_2 = 3.64$), which means the scores are spread out over a relatively similar range in the females' and males' distributions.

The foreign language streams examinees' scores' descriptive statistics, on the other hand, are considered to be the most important since they were quite different from the previous ones and opened doors, together with the corresponding inferential statistics, for

answering the research questions and ending up with insightful conclusions. The first finding was that the scores mean for males and females were very close to each other in both exams (BAC 2014: $\mu_1 = 10.02 / \mu_2 = 10.33$ & BAC 2019: $\mu_1 = 9.17 / \mu_2 = 9.42$). The same can be said about the median in BAC 2014, where: $MDN_1 = MDN_2 = 10.00$. In BAC 2019, the median of the male population was higher than that of the females and this is a unique finding ($MDN_1 = 9.50 / MDN_2 = 9.00$). This leads to considering the skeweness values for the males' distributions in tables 5.11 and 5.12, which are both negative. This implies that the left tail is longer compared to the right tail i.e. the number of scores above the mean is greater than that below it, which means that this male population have achieved better than all the other populations in terms of distribution of scores. The female population scores' distributions were positively skewed and approximately symmetric.

The results of the two sample z-tests indicated that for the scientific stream examinees, there is a significant difference between male and female score means in both exams. The p -values (BAC 2014: $p = 0.000339$ & BAC 2019: $p = 1.42E-14$) were both lower than α , where $\alpha = 0.05$. Therefore, this research's null hypothesis (H_0) was once more rejected and the second research (alternative) hypothesis was confirmed. Yet, the calculation of the effect size to check the magnitude of the difference between the two groups has shown that it is very small as both values (BAC 2014: $d = 0.07$, BAC 2019: $d = 0.04$) were below Cohen's (1988) conventions of small effect ($d = 0.2$). This refers that the difference can be neglected.

However, the z-tests results for the foreign languages stream have led to different conclusions. Since the p -values for the two exams (BAC 2014: $p = 0.610099$, BAC 2019: $p = 0.453693$) were both higher than α , where $\alpha = 0.05$, the null hypothesis this time was confirmed, which means that the difference between the male and female populations is not big enough to be statistically significant. This finding, together with the results of the

descriptive statistics, especially for the male population, implies that there should be something specific related to the status of the English language exam and the English language as a subject within this particular stream that led to these exceptional results compared to the other exams. This would provide an answer to the third research question that would be discussed in the following point together with the other questions.

Summary of the Discussions

In this section, and based on all the findings discussed above, the three research questions would be answered and the research hypotheses would be discussed to either be confirmed or rejected. Links are made between the first three chapters, which are theoretical in nature and constitute a solid background for conducting the research, from the one side, and the findings of textbooks' and exam papers' analysis and evaluation, the exams' specifications, and the statistical data analysis from the other side.

Research question N°01: Are the testing and the instructional English language materials (i.e. exams and textbooks) designed for Algerian middle and secondary school learners, sitting for the BEM and BAC exams, homogeneous in terms of topics and task/activity types?

In other words, this is aimed to question the *validity* (**See chapter two**) of the BEM and BAC English language exams. The English language textbooks used by middle and secondary school teachers and learners in the final year of these cycles provided sufficient data about the adopted syllabus type (both textbooks have a *Multi-syllabus syllabus*), and the topics and task/activity types included within these instructional materials. At the same time, both textbooks were found to follow the guidelines and the course objectives determined in the curricula of the English language for both school years. In addition to that, the analysed BEM and BAC exam papers have strictly stuck to the typologies of activities

set by the ministry of national education in Algeria, which is the organism responsible for the whole educational and testing policy in the country. Therefore, and based on what has been discussed earlier, and following Palmer's (1981) words on *research about test validity* saying that in general, it makes "...investigations into the extent to which a test measures what it is supposed to measure" (p. 135), the first null hypothesis stating that: *There is a high degree of homogeneity between the BEM and BAC English language exam papers and the instructional textbooks used by Algerian learners in terms of topics and task/activity types cannot be rejected* and the exams have achieved *test validity*.

Research Question N°02: Is there a gender gap in Algerian middle and secondary school learners' academic achievements in BEM and BAC exams? If yes, to which extent?

In order to answer this question, it was necessary to study and compare male and female examinees' scores at two different points in time and not be satisfied by examining them based on the results of one single exam as this would give lower credibility to results. In a way, this has its background in the theoretical background related to *test reliability*, which is often considered as consistency in measurement (Bachman & Palmer, 1996). The descriptive and inferential statistics results have shown that the examinees' scores, scores means, scores distributions and differences between the means, were very similar in the BEM and BAC exams in the years 2014 and 2019. This can tell a lot about these exams' reliability. Moreover, these are state, state-wide and blind peer-rated exams. This gives more value and credibility to the research results.

The findings, as was previously discussed, have revealed that apart from the foreign languages stream examinees, whose results have shown that the difference between the male and female populations is not big enough to be statistically significant, all the other analyses

demonstrated that the difference is statistically significant. However, the calculation of Cohen's *d* coefficient of effect size proved that the magnitude of the difference is very small (lower than 0.2), which makes it negligible. Hence, the answer to the above question is *NO* and taking the results, on the whole, the second null hypothesis, which states that: *There is no difference between male and female learners' means in the BEM and BAC English language exams cannot be rejected.*

Research Question N°03: Based on the above questions, can gender be considered a primary indicator of academic achievements?

Taking into account the fact that the analysed BEM and BAC exams are *valid* and *reliable*, and that their results were studied at two different points in time giving the same results, we can make some inferences by referring to some of the exams' specifications dealt with in *Chapter Four* and male and female populations' and samples' results summarised in this chapter, and previously analysed and discussed in *Chapter Five*.

As it was explained in Chapter Four, English in the BEM exam is given two (02) as a coefficient, which makes it a secondary subject compared to other subjects like Arabic (coefficient: 05), maths (coefficient: 04), History & Geography (coefficient: 03) or even French (coefficient: 03). The same status is given to English in the BAC exam for the scientific stream i.e. the coefficient of the English language exam is two (02), lagging behind other fundamental subjects like Natural sciences (coefficient: 06), Maths (coefficient: 05), Physics (coefficient: 05), and Arabic language (coefficient: 03). Although the effect size of the difference between male and female populations was very small, their difference was statistically significant, and according to the analyses of the descriptive statistics, females had better results.

Nevertheless, the findings of the analysis of the foreign languages stream male and female examinees' score means gave other insights as the difference was not big enough to be statistically significant. Moreover, males, according to the descriptive statistics, especially in terms of the score distributions, had a better performance than females, and than all the other populations' performances in the other discussed exams' results. The fact that English in this particular stream is given the coefficient of (5) together with Arabic and French being at the top of the list makes it a fundamental subject and one that needs to be taken seriously by the learners seeking success in the Baccalaureate exam.

Based on the above analysis and discussion, it seems that when English is given high status, it does not appear that there would be any statistically significant difference between males and females' means but males would even show a better performance than females as it was the case for the foreign languages stream learners. Therefore, the answer to the aforementioned question is also: *NO, gender is not a primary indicator of learners' academic achievements.*

6.3. Recommendations and Implications

6.3.1. To the Policy Makers and Curriculum Developers

The *gender gap* in educational outcomes has long been a topic of interest and research in the academic field worldwide. This so-called 'phenomenon' has, also, always caught the attention of educational specialists, the media and a very wide range of society since learners' academic achievements are directly related to and affect in one way or another the family being the nucleus of the society. Therefore, very special attention needs to be given by the policymakers i.e. the political authorities, and the curriculum developers as well for the sake of giving equal chances for males and females to achieve the best based on their

individual capacities and not because the educational system is biased - explicitly or implicitly- towards one category or the other.

The National Education Orientation Law (Ministry of Education, 2008, p.39) includes a band stating that one of the objectives of the Algerian educational system is to *ensure an equal schooling rate between girls and boys at all levels of education* is very encouraging. However, the question that arises is **how and with what measures?** Moreover, maybe **what is more essential than ensuring equal schooling rates is to ensure equal chances of success for everyone and to enhance the status of the school and education in general within the society.** When this is achieved, learners would compete starting all from the same line, and gender rates of success would become a secondary variable and not a primary concern for the educational authorities as the competition would be on an individual basis regardless of their gender.

The majority of the research studies referred to in *Chapter One* have concluded that males and females show fewer differences than similarities. This is manifested mainly in terms of cognitive abilities such as intelligence, perception, verbal and quantitative abilities. The results of the present study correlate largely with these findings. It was found that when English as a subject is valued intrinsically (i.e. being the learner's choice to study foreign languages), as well as extrinsically (i.e. it is given a high coefficient that raises its status and contribution to pass the exam), the result was that there was no significant difference between the score means of males and females. This leads to thinking about the psychological aspect, in which the two genders show some differences according to the studies previously referred to in *Chapter One*. This aspect needs to be given much importance by the educational authorities through raising the syllabus designers, test designers and administrators, as well as teachers' (during both pre-service and in-service training) awareness about these differences and making clear guidelines and relevant training

content to deal with them for the sake of avoiding bias as much as possible. Similar importance should be given, especially in the teachers' training programmes designed by the educational authorities often by cooperating with the ministry of higher education and scientific research, to the biological differences and the physiological developments of males and females as they are closely linked to the psychological state of individuals at different age stages and contribute in enhancing the psychological differences between them. All this has, for sure, to be controlled and inspected in the field to ensure that every component of the educational system is completing the attributed tasks in the best manner.

6.3.2. To the Syllabi and Textbook Designers

The textbook, which is the application and reflection of the syllabus, is of paramount importance in the teaching/learning process. Syllabus and textbook designers need to take gender issues into account during the design process. The elements of textbook analysis and evaluation adopted in this research (i.e. topics and activity types) and the results of this analysis would help, during the design of future textbooks, in making a better selection, of bias-free topics in addition to a balanced set of activities and tasks dealing with:

- all four language skills (i.e. *reading, writing, speaking, and listening*),
- all the language components (*vocabulary, grammar, pronunciation, morphology and discourse*),
- and that requires the learners to engage in them through a balanced use of *lower-order (knowledge, comprehension, and application)* and *higher-order (analysis, synthesis and evaluation)* cognitive abilities.

The designers of the 'syllabus' that is considered as the overall organizing principle for what is to be taught and learned (McDonough et al., 2013), apart from the classroom practices that are believed by many researchers to be part of *methodology*, and who are often

the textbooks' designers, have to consider gender differences and similarities, especially the psychological, social and cognitive ones in the design process. Gender-stereotypical roles, practices, language use, and behaviours should be avoided as much as possible in the syllabus and textbooks' contents. These can influence learners' attitudes towards the subject and as a result, it would affect their performance on tests and exams, which is not the desired consequence, especially when the effect is negative.

6.3.3. To the Exams' Designers

In addition to achieving validity, reliability, practicality and washback, one major criterion of a *good test* is *equivalency*. This is achieved, according to Mihai (2010) if a test is directly based on curricular standards or instructional activities. The results of this research have shown that the BEM and BAC exams, which are state-wide examinations, fit all the cardinal criteria above mentioned. In terms of *equivalency*, the exams' designers have shown that they strictly stick to the BEM and BAC exams guidelines set by the ministry of national education and manifested in policy documents such as the **BEM Guide for the English Language Examination (October, 2018)**, and the **BAC Guide for the English Language Examination (October, 2017)**, especially concerning the typology of activities. This could be also noticed in the choice of topics treated in the exam papers that were inspired by the English language syllabi and textbooks for the fourth-year middle school and third-year secondary schools.

The textbooks' analysis has shown that all language skills are present; however, both BEM and BAC exams, contrary to most international and widely recognised tests and examinations, tend to test only the reading and writing skills in addition to the different language components. The fact that these state-wide examinations do not consider assessing learners' listening and speaking abilities to be essential in determining one's proficiency in

a language represents a serious disadvantage and lessens the extent to which we can consider a learner's score on that exam as a true indicator of his/her overall language abilities. Language exam designers need to report this to the educational authorities since they are supposed to be the experts from whom the decisions concerning the testing policy are inspired.

6.4.Limitations of the Study

The political and educational reforms and changes, especially the move from the first-generation to the second-generation curriculum and the troubles related to exam-questions' leaks in the period in which the present research was conducted disrupted it a bit at its rate of advance. Moreover, the fact that the exam papers are peer- and blind-corrected, it was difficult to get the corrected answer sheets, and even when this was possible, the use of a code instead of the examinees' names on them (blind correction) made the task harder to identify whether it is a male or a female examinee's paper. This was intended to extend the research's work to examine male and female examinees' performances on individual exam activities and compare some samples' marks. That would give better insights into male and female learners' achievements and the link with gender differences would have been more informative.

6.5.Suggestions for Further Research

Since this research focused on populations belonging to the same part of the country (Bordj Bou Arreridj), the statistical analyses adopted in this research can be extended to other parts of Algeria or even be used on a nationwide scale in order to reach more generalizable conclusions about the reality of the gender gap in Algerian learners academic achievements in English and possibly in other school subjects. More research can be conducted also in terms of male and female learners' performances on different test items

and examining the possible relationship with gender differences in cognitive abilities for example. This would better explain differences between males and females not only based on a test's final scores but in-tests items' separate scores.

Conclusion

In this chapter, the three research questions were answered and the research hypotheses were also tested based on the findings and discussed presented in the previous chapters. It has been found that all the questions raised in this research have received negative answers. The conclusion is that the BAC and BEM exams are valid ones as they show a very high degree of homogeneity with the instructional textbooks used in the Algerian middle and secondary schools textbooks. Moreover, there seems to be no significant difference between male and female populations' means of scores in the BEM and BAC exams in Algeria. As a result, it was concluded that gender could not be considered a primary indicator of academic achievements since the same conditions lead to the same results in all the analysed learners' performances. Therefore, other factors related to the subject itself and the status given to it and its exam within the overall educational system may play a greater role in influencing academic achievements than the gender variable. Hence, some implications and recommendations were suggested for policymakers and curriculum, syllabus, textbook, and exam designers.

General Conclusion

Nowadays, and in many parts of the world, there is a wide reference to what is considered to be a *reverse gender gap* meaning that females tend to perform better than males at different school levels and on various school subjects. This is supported by the statistics found in the reports about males' dropout from school at an early age. Hence, the primary aim of this study was to find out if these claims are true for Algerian learners studying English as part of the overall curriculum governing the middle and secondary schools' teaching/learning. In order to achieve this aim, the research adopted a descriptive cross-sectional approach by:

- First, analysing the instructional materials (On the Move & New Prospects textbooks) as well as the testing materials (English language BEM and BAC exam papers) since they are basic parts of the curriculum.
- And then, middle and secondary school learners' achievements were analysed making use of descriptive and inferential statistics.

On the basis of the findings of these research tools and procedures, it was concluded that the analysed instructional and testing materials show a high degree of homogeneity and; therefore, there is little chance of dealing, in the BEM and BAC exams, with topics or facing activities different from those encountered in the course books. In addition to that, there was little evidence that there is statistically significant difference between male and female examinees' score means and even when there was a statistical evidence, the magnitude of the difference reflected by Cohen's *d* coefficient of effect size was very small to the extent that it could be neglected. From what precedes, it was concluded that gender cannot be taken as an indicator of achievement and a generalization about which sex performs better than the other is not possible. Saying that females perform better than

males or vice versa may be misleading. A better description of the finding would be that *some females perform better than some males and vice versa*.

Comparing the results of the statistical analysis of foreign languages stream BAC examinees with their counterparts from the scientific stream and the BEM exam as well revealed that males' performance was noticeable as their distribution of scores was skewed to the left which was a unique finding in this research. Added to that is the fact that the results of the two sample z-tests showed that the mean of male examinees is considered to be equal to the mean of females. This led to scrutinizing the exam characteristics and what makes it different from the other treated exams. The answer was found in the coefficient given to English that is the highest one (05) equally with Arabic and French. Compared to the BEM exam and the BAC exam for the scientific stream, in which English has a coefficient of two (02), this led to the belief that the learners' attitude towards English being a fundamental subject and good performance in its exam helps a lot in achieving success in the overall exam makes the difference.

In light of the obtained results, some implications and recommendations were suggested and addressed to:

- First, policymakers and curriculum designers in order to think about learners not on the basis of their gender, but rather as being individuals, who are more similar than different. Moreover, their policies need to be built on the belief that their educational outcomes are not determined by gender, but by their attitudes towards learning in general and the importance they give to one subject or another based on the extent to which a good achievement in it would contribute to success.
- Second, syllabus and textbook designers, need to design *unbiased* materials by avoiding stereotypical ideas about males' and females' roles for example in the topics treated in

the textbooks. Moreover, they need to ensure that the activities and tasks they include are balanced in terms of the language skills, the language components, the targeted cognitive abilities and skills in addition to the nature of the language input and output.

- Finally, last but not least, exam designers, who are responsible for making those exams fit the criteria of a *good* test i.e. the exams need to be *valid*, *reliable*, *practical* and achieve *equivalency* and *washback*. It was found that the BEM and BAC exams are congruent with all these criteria. Yet, the black spot in all that white paper is the absence of a listening and speaking section. This has to be reconsidered since, language standardized tests, which are often taken as a reference in language assessment, do test language learners' listening and speaking abilities since they represent the aural/oral parts of the language

On the basis of the present study, further research can be conducted to scrutinise male and female learners' performances on different item types in either exam (BEM or BAC) and examine their relationship with males and females' cognitive abilities and differences in the brain for example.

REFERENCES

- Allen, J.P.B. (1984). General-Purpose Language Teaching: A Variable Focus Approach. In Brumfit C.J. (Ed.), *General English Syllabus Design* (pp.61-74). Pergamon Press Ltd. and the British Council.
- Alderson, J. C. & Hughes A. (Eds.) (1981). *ELT documents: Issues in Language Testing*. London: The British Council.
- American Psychological Association. (2020). Publication manual of the American Psychological Association 2020: the official guide to APA style (7th ed.). American Psychological Association.
- Anderson, L.W. & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York: Longman.
- Arab, S.A. & Riche, B. (2006). *On the Move*. Algiers: The National Authority for School Publications.
- Arab, S.A. et al. (n.d.). *New Prospects (Revised Edition)*. Algiers: The National Authority for School Publications.
- Bachman, L.F. (1990). *Fundamental Considerations in Language Testing*. Oxford University Press: Oxford.
- Bachman, L.F. & Palmer, A.S. (1996). *Language Testing in Practice: Designing and Developing Useful Language Tests*. Oxford University Press: Oxford.
- Becker, W.C. (1986). *Applied Psychology for Teachers: A Behavioral Cognitive Approach*. Chicago: Science Research Associates, Inc.

- Black, P. & William, D. (1998). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice*, 5 (1), 7- 74.
- Bloom, B. S. (1956). *Taxonomy of Educational Objectives, Handbook I. The Cognitive Domain*. New York: David McKay Co Inc.
- Brannon L. (2017). *Gender – Psychological Perspectives (7th ed.)*. New York: Taylor & Francis.
- Breen, M.P. (1984). Process Syllabuses for the Language Classroom. In C.J. Brumfit (Ed.), *GENERAL ENGLISH SYLLABUS DESIGN: Curriculum and syllabus design for the general English classroom* (pp. 47-60). Pergamon Press Ltd. and the British Council.
- Breen, M. (1987). Contemporary Paradigms in Syllabus Design. Part I. *Language Teaching*, 20(2), 81-92. DOI: 10.1017/S0261444800004365
- Brown, H.D. (2004). *Language Assessment: Principles and Classroom Practises*. Pearson Education, Inc: New York.
- Brown, J.D. (1995). *The Elements of Language Curriculum: A Systematic Approach to Program Development*. Boston: Heinle & Heinle Publishers.
- Brown, J.D. (1996). *Testing in Language Programs*. New Jersey: PRENTICE HALL REGENTS.
- Brumfit C.J. (Ed.) (1984). *GENERAL ENGLISH SYLLABUS DESIGN: Curriculum and syllabus design for the general English classroom*. Pergamon Press Ltd. and the British Council.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences (2nd Ed.)*. LAWRENCE ERLBAUM ASSOCIATES, PUBLISHERS.

- Cohen, R. J. & Swerdlik, M. (2009). *Psychological Testing and Assessment: An Introduction to Tests & Measurement*. McGraw-Hill Higher Education.
- Cole, M. (Ed.) (2003). *Education, Equality and Human Rights*. Taylor & Francis e-Library.
- Cunningsworth, A. (1995). *Choosing Your Coursebook*. Oxford: Macmillan Education.
- Daniels, L. & Minot, N. (2020). *An Introduction to Statistics and Data Analysis Using Stata®*. SAGE Publications, Inc.
- Ebel, R.L. and Frisbie, D.A. (1991). *Essentials of Educational Measurement 5th Edition*. Prentice-Hall, Englewood Cliffs.
- Ellis, R. (2011). Macro- and micro-evaluations of task-based teaching. In B. Tomlinson (Ed.), *Materials Development in Language Teaching* (2nd ed., pp. 212-235). Cambridge: Cambridge University Press.
- Everhard, C.J. & Murphy L. (Eds.) (2015). *Assessment and Autonomy in Language Learning*. PALGRAVE MACMILLAN.
- Genesee, F. & Upshur J.A. (1996). *Classroom-based Evaluation in Second Language Education*. Cambridge: Cambridge University Press.
- Gullo, D.F. (2005). *Understanding Assessment and Evaluation in Early Childhood Education*. New York: Teachers College Press.
- Haas, A. (1979). Male and Female Spoken Language Differences: Stereotypes and Evidence. *Psychological Bulletin*, 86 (3), 615-626.
- Halpern, D. F. (2004). A Cognitive-Process Taxonomy for Sex Differences in Cognitive Abilities. *Current Directions in Psychological Science*, 13 (4), 135-139.
<http://dx.doi.org/10.1111/j.0963-7214.2004.00292.x>
- Halpern, D.F. (2012). *Sex Differences in Cognitive Abilities (4th ed.)*. New York: Taylor & Francis.

- Harmer, J. (2007). *The Practice of English Language Teaching* (4th ed.). Essex: Pearson Education Limited.
- Holmes, M. (2007). *What is Gender? Sociological Approaches*. London: SAGE Publications Ltd.
- Hughes, A. (1989). *Testing for Language Teachers*. Cambridge: Cambridge University Press.
- Hyde, J.S. & McKinley, N.M. (1997), Gender Differences in Cognition: Results from Meta-Analyses. In P.J. Caplan, M. Crawford, J.S. Hyde & J.T.E. Richardson, *GENDER DIFFERENCES IN HUMAN COGNITION* (pp.30-51). New York: Oxford University Press, Inc.
- Hyde, J.S. (2005). The Gender Similarities Hypothesis. *American Psychologist*, 60 (6), 581–592. DOI: 10.1037/0003-066X.60.6.581.
- Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger Publishers.
- Jones, C.H. (2005). *Assessment for Learning*. Learning and Skills Development Agency, UK.
- Krahnke, K. (1987). *Approaches to Syllabus Design for Foreign Language Teaching*. Englewood Cliffs, NJ: PRENTICE-HALL, INC.
- Kubiszyn, T. & Borich, G. (2013). *EDUCATIONAL TESTING AND MEASUREMENT: Classroom Application and Practice* (10th ed.). New Jersey: John Wiley & Sons, Inc.
- Lane, D. (2022). *Introductory Statistics*. LibreTexts.
- Lane, D.M. (n.d.). *Online Statistics Education: A Multimedia Course of Study*. Retrieved from: <http://onlinestatbook.com/>. Rice University.
- Littlejohn, A. (2011). The analysis of language teaching materials: inside the Trojan Horse. In B. Tomlinson (ed.), *Materials Development in Language Teaching* (2nd ed., pp. 179-211). Cambridge: Cambridge University Press.

- Marzano, R.J. & Kendall, J.S. (2007). *The New Taxonomy of Educational Objectives* (2nd ed.). California: Corwin Press.
- McDonough, J., Shaw, C. & Masuhara, H. (2013). *Materials and Methods in ELT: A Teacher's Guide* (3rd ed.). West Sussex: John Wiley & Sons, Inc.
- McGrath I. (2002). *Materials Evaluation and Design for Language Teaching*. Edinburgh: Edinburgh University Press Ltd.
- Messick, S. (1994). *Validity of Psychological Assessment: Validation of Inferences from Persons' Responses and Performances as Scientific Inquiry into Score Meaning*. Research Report RR-94-45. Princeton, N.J.: Educational Testing Service,
- Meyer, E.J. (2010). *Gender and Sexual Diversity in Schools*. London New York: Springer Science + Business Media.
- Mihai, F.M. (2010). *Assessing English Language Learners in the Content Area: A Research-into-Practice Guide for Educators*. Michigan: The University of Michigan Press.
- Ministry of National Education. (2005, January). PROGRAMME D'ANGLAIS DEUXIEME LANGUE ETRANGERE (première année secondaire) [Curriculum of English as a Second Foreign Language (first year secondary school)]. Ministry of National Education of Algeria.
- Ministry of National Education. (2007). Curriculum of English: Third Year Secondary School. Ministry of National Education of Algeria.
- Ministry of National Education. (2008, January 23). Al-quanoun Al-twjihi Li Tarbiya [The National Education Orientation Law], N°08-04. Ministry of National Education of Algeria.
- Ministry of National Education. (2009, March). Al-marjiiya Al-aama Li Al-manahij [The General Reference for the Curricula]. Ministry of National Education of Algeria.
- Ministry of National Education. (2016). Al-baramij Al-dirassiya Li Tawr Al-awal Min Marhalat Al-taalim Al-moutawassit [The Educational Programs for the First Stage of Middle School Education]. Ministry of National Education of Algeria.

- Ministry of National Education. (2017, October). *Dalil Binaa Ikhtibar Madat Al-Lougha Al-Ingliziya fi Chahadat Al-Baccalaureat* [The English Language Test-Design Guide for the BAC Exam]. Ministry of National Education of Algeria.
- Ministry of National Education. (2018, April). *TEACHER'S GUIDE: Middle School Key stage 3/Year Four*. Boumerdes: Ministry of National Education of Algeria.
- Ministry of National Education. (2018, October). *Dalil Binaa Ikhtibar Madat Al-Lougha Al-Ingliziya fi Chahadat Al-Taalim Al-Moutawassit* [The English Language Test-Design Guide for the BEM Exam]. Ministry of National Education of Algeria.
- Nachmias, C.F. & Gerrero, A.L. (2018). *Social Statistics for a Diverse Society (8th ed.)*. California: SAGE Publications, Inc.
- Nunan, D. (1988). *Syllabus Design*. Oxford: Oxford University Press.
- Nunan, D. (1989). *Designing Tasks for the Communicative Classroom*. Cambridge: Cambridge University Press.
- Nunan, D. (1999). *Second Language Teaching & Learning*. Boston: Heinle & Heinle Publishers.
- Nunan, D. (2004). *Task-Based Language Teaching*. Cambridge: Cambridge University Press.
- Nunan, D. (2015). *TEACHING ENGLISH TO SPEAKERS OF OTHER LANGUAGES: An Introduction*. New York: Routledge.
- OpenStax College, *Introductory Statistics* OpenStax College. 13 August 2015. <http://cnx.org/content/col11776/latest>.
- Palmer, A.S. (1981). BASIC CONCERNS IN TEST VALIDATION. In J. C. Alderson, & A. Hughes (Eds.), *ELT documents: Issues in Language Testing* (pp. 135-151). London: The British Council.
- Peters, R.S. (Ed.) (2010). *The Concept of Education*. Oxon: Taylor & Francis e-Library.
- Prabhu, N.S. (1987). *Second Language Pedagogy*. Oxford: Oxford University Press.

- Richards, J.C. (1990). *The Language Teaching Matrix*. Cambridge: Cambridge University Press.
- Richards, J.C. (2001). *Curriculum Development in Language Teaching*. Cambridge: Cambridge University Press.
- Richardson, J.T.E. (1997). Introduction to the Study of Gender Differences in Cognition. In P.J. Caplan, M. Crawford, J.S. Hyde & J.T.E. Richardson, *GENDER DIFFERENCES IN HUMAN COGNITION*. New York: Oxford University Press, Inc.
- Richardson, S.S. (2012). Sexing the X: How the X Became the “Female Chromosome”. *Signs*, 37 (4), Sex: A Thematic Issue (Summer 2012), 909-933.
<http://www.jstor.org/stable/10.1086/664477>
- Rhoads, S.E. (2004). *Taking Sex Differences Seriously*. San Francisco: Encounter Books.
- Samuda, V. & Bygate, M. (2008). *Tasks in Second Language Learning*. New York: PALGRAVE MACMILLAN
- Statt, D.A. (2003). *The Concise Dictionary of Psychology (3rd ed.)*. New York: Taylor & Francis e-Library.
- Tomlinson, B. (ed.) (2011). *Materials Development in Language Teaching (2nd ed.)*. Cambridge: Cambridge University Press.
- Ullman, M.T., Miranda, R.A., & Travers, M.L. (2008). Sex Differences in the Neurocognition of Language. In J.B. Becker, K.J. Berkley, N. Geary, E.H. Hampson, J.P. Herman, & E.H. Young (Eds.), *Sex Differences in the Brain from Genes to Behaviour* (pp. 291- 310). Oxford: Oxford University Press, Inc.
- Widdowson, H.G. (1984). Educational and Pedagogic Factors in Syllabus Design. In C.J. Brumfit (Ed.) (1984). *GENERAL ENGLISH SYLLABUS DESIGN: Curriculum and syllabus design for the general English classroom* (pp. 23-28). Pergamon Press Ltd. and the British Council.

Wilkins, D.A. (1976). *Notional Syllabuses. A Taxonomy and its Relevance to Foreign Language Curriculum Development*. Oxford: Oxford University Press.

Winch, C. & Gingell, J. (2008). *Philosophy of Education: The Key Concept (2nd ed.)*. Routledge.

Appendix A

A Schedule for Analysing Tasks (Littlejohn, 2011, p.191)

Task Analysis Sheet								
Task number:								
I. What is the learner expected to do?								
A. TURN-TAKE								
Initiate								
Scripted response								
Not required								
B. FOCUS								
Language system (rules or form)								
Meaning								
Meaning/system/form relationship								
C. MENTAL OPERATION								
<i>[detailed according to what is found in the materials]</i>								
II. WHO WITH?								
<i>[detailed according to what is found in the materials]</i>								
III. WITH WHAT CONTENT?								
A. INPUT TO LEARNERS								
Form								
Source								
Nature								
B. OUTPUT FROM LEARNERS								
Form								
Source								
Nature								

Appendix B: Sample Tasks' Analysis Sheet (Adapted from Littlejohn (2011, p191))

Task analysis sheet	File 01 (Listen and Consider)											File 01 (Practice)					
Task number	1	2	3	4	5	6	7	8	9	10	11	1	2	3	4	5	6
1- WHAT IS THE LEARNER EXPECTED TO DO?																	
A- TURN TAKE																	
Initiate																	
Scripted response																	
Not required																	
FOCUS: LANGUAGE COMPONENTS AND SKILLS																	
Vocabulary																	
Grammar																	
Pronunciation																	
Morphology																	
Discourse																	
Listening																	
Speaking																	
Reading																	
Writing																	
COGNITIVE/THINKING SKILL																	
Knowledge																	
Comprehension																	
Application																	
Analysis																	
Synthesis																	
Evaluation																	
2- WHO WITH?																	
Learners individually simultaneously																	
Pair work / group work																	
Learner to whole class																	
Learner individually outside the class																	
3- WITH WHAT CONTENT?																	
A- INPUT TO LEARNER: FORM																	
Extended discourse: written																	
Extended discourse: aural																	
Words/phrases/sentences: written																	
Words/phrases/sentences: aural																	
Graphic																	
B- EXPECTED OUTPUT FROM LEARNERS: FORM																	
Extended discourse: written																	
Extended discourse: aural																	
Words/phrases/sentences: written																	
Words/phrases/sentences: aural																	
Graphic																	
Other (e.g. ↗ , `,...)																	

Appendix C

'On the Move' Coursebook Adaptation Chart

Course book Adaptation – April 2018

		PROJECTS	TOPICS	ON THE MOVE COURSE BOOK	
The 1st Term	Sequence.1	A moudjahid portrait	ME My personality and experiences: -Personality features -School and life experiences (embarrassing and positive ones)	File 04 : P.102/103/ File 06 : P.143/144/146/147/148/149	Assessment & Remediation Where do we stand now? /Learning log (to be exploited by the teacher)
	Sequence.2		ME My prospects and initiatives -Projects and initiatives -Career/ education	File 02 : P.46/47 File 03 : P.50/68 /67/68/69/70/71/72/73/77 P.82/83/84	
The 2nd Term	Sequence.3	Designing Charity advertisement leaflet	My World HEALTHY FOOD -Healthy food and table manners -Family advice Citizenship: Commitment with charity/ clubs/ helping hand/ respect of others	File 01 : ALL PAGES File 03: Words & sounds (Diphthongs) P99	Teachers can select the appropriate statements to elaborate the assessment grid
The 3rd Term	Sequence.4	Travelling on line	The World Online Travelling : Forums of discussion Networks and Web sites on: - Food and dishes around the world -Landmarks and monuments -Outstanding figures in history, literature and arts	File 03 P.76/78/79/80/81 File 04 P.101/107 /109/110/128/129/ P.96/97/104/105/106	

Appendix D

Structure and Typology of the BEM Exam Activities

1- Structure of the Exam

Parts	Sections	Activities	Scoring scale
① (14 pts)	A/ Reading Comprehension	✓ 2 activities related to comprehension	<u>Act.1</u> 03points <u>Act.2</u> 02points
		✓ 1 activity related to Lexis	<u>02 pts</u>
	B/ Mastery of Language	✓ 1 activity related to mechanics or morphology.	<u>02 pts</u>
		✓ 1 activity related to syntax	<u>03 pts</u>
		✓ 1 activity related to pronunciation	<u>02 pts</u>
② (06 pts)	Situation of Integration The candidates are supposed to write a communicative passage. These types of tasks may help: ✓ Letter-writing ✓ E-mail ✓ Dialogue ✓ Report ✓ Article ✓ Social media ✓		(06 pts)

2- Activity types

Parts		Tasks	Instructions
Part one: A- Reading comprehe nsion	Comprehension	Matching	-Match each title/heading/idea with the corresponding paragraph. - Match the dates with the corresponding events. - Match the answers in column B with the corresponding questions in column A.
		Information Transfer	-Complete the table /diagram/plan/chart/form... with information <u>from the text</u>. -Complete the table with the events that correspond to the dates.
		MCQ	-Choose (a, b, c or d) to complete the following sentences. -Which sentence best summarizes the text (provide four sentences). -Choose the most appropriate title (provide four titles). -The text is about : a)...../ b)...../ c)...../ d).....
		Auxiliary Questions.	-Answer the following questions.(yes/no questions)
		Open Ended Questions	-Answer the following questions (inference questions). - Eg: What is ...?
		Listing	-Find in the text four similarities/differences between (people, places, ...) -Find in the text two similarities and two differences between.....
		True/False /Not mentioned	-Read the following statements and write "True"/ "False" or "Not mentioned"
		Identification	-What do the underlined words in the text refer to? - In which paragraph is it mentioned that?
	Lexis	Matching	-Match each word with its appropriate definition /meaning. -Match the words in column a with their equivalents/opposites in column B. -Match each verb with its appropriate definition. <u>Note:</u> Accept the answer if the learners match numbers with letters .
		MCQ	-What do these words mean?- Choose the right word (provide four meanings/synonyms for each word). -Eg: clever=..... a)stupid b)ignorant c)intelligent d)witty
		Finding	-Find in the text words/ expressions that have the following definitions/meanings. -Find in the text words that are closest in meaning to the following. -Find in the text words that are opposite in meaning to the following.
			-Rewrite the following passage or sentences using the suitable

Part one: B Mastery of Language	Morphology / Mechanics / Syntax .	Transformation tense, -Turn the following sentences /passage into the negative /passive/active/plural/singular/..... -Make a noun /verb/adjective/adverb from each of the following words: -Choose the correct prefix to form the opposites of these words: -Add the correct prefix to form the opposite adjective. -Note: Provide the adjectives (e.g.: kind – polite – patient – regular -sensitive etc.....) <table><tr><td>un</td><td>in</td><td>in</td><td>ir</td></tr><tr><td>1</td><td>1 2</td><td>1</td><td>1</td></tr></table> -Write the noun forms of the following words, using a suffix. <table><tr><td>Act</td><td>.....</td></tr><tr><td>Science</td><td>.....</td></tr><tr><td>Write</td><td>.....</td></tr><tr><td>Music</td><td>.....</td></tr></table> -Rewrite the following sentences using the adverb corresponding to the adjective. eg. : My brother is a quick worker. My brother works quickly.	un	in	in	ir	1	1 2	1	1	Act		Science		Write		Music		Editing -Spot the mistakes and write the corrected sentences (or passage) on your answer sheet. -Correct the underlined mistakes. Eg.: My friend celebrated his birthday. He invited me four lunch when we are in skikda, yestrday. Note: The passage must be related to the text . -Supply the punctuation and the capital letters where necessary.	Completion -Fill in the gaps with the suitable words. They can either be provided or not. -Complete the following conversation with the right questions. -Complete the following sentences with the correct relative pronoun.(in context) -Complete the following sentences using the comparative form / superlative form of the adjective in brackets. (in context)	Combining -Combine the following phrases (or-sentences) using connectors, linking words ...	Sound Recognition -In each list, pick out the word with a different vowel sound. (Four words in each list). Example: date –late –mat-lane Note: Give 4 words containing the same vowel letter among which one has a different vowel sound. -Write the following words in the right box according to their consonant sound : "couple/slice/cute/rice" <table><tr><td>/S/</td><td>/K/</td></tr><tr><td>1.</td><td>1.</td></tr><tr><td>2.</td><td>2.</td></tr></table>	/S/	/K/	1.	1.	2.	2.
		un	in	in	ir																							
		1	1 2	1	1																							
		Act																										
		Science																										
Write																												
Music																												
/S/	/K/																											
1.	1.																											
2.	2.																											
Pronunciation																												
		Matching -Match the words that have the same rhyme. <table><tr><td>A</td><td>B</td></tr><tr><td>1-fair</td><td>a-late</td></tr><tr><td>2-make</td><td>b-meet</td></tr><tr><td>3-tough</td><td>c-enough</td></tr><tr><td>4-meat</td><td>d-pair</td></tr></table>	A	B	1-fair	a-late	2-make	b-meet	3-tough	c-enough	4-meat	d-pair																
A	B																											
1-fair	a-late																											
2-make	b-meet																											
3-tough	c-enough																											
4-meat	d-pair																											

Source: The English Language Test-Design Guide for the BEM Exam (October, 2018)

Appendix E

Typology of the BAC Exam Activities

1- Comprehension

Type of activities			Number of activities
1	OR	<u>Type of text</u> *The text is a... a) letter b) conversation...	(Not more than 04 statements)
		<u>T / F statements</u> *Are the following statements true or false? Write T or F next to the letter corresponding to the statement. * Are the following statements true or false? Write T or F next to the letter corresponding to the statement and correct the false one(s).	
		<u>MCQ</u> *Write the letter that corresponds to the right answer.	
2	OR	<u>Paragraph identification</u> *In which paragraph is it mentioned that...? OR *In which paragraph	Identifying 2 Paragraphs
		<u>Sentence order</u> * Read the text and put the following ideas in the order they	3 to 4 ideas
		<u>Table completion</u> *Fill in the following table with information from the text.	4 items to be included in the table
3		<u>Answering comprehension questions</u> *Answer the following questions according to the text.	* include ONE <u>inference</u> question 3 to 4 questions
4		<u>Cohesive markers</u> *Who or what do the underlined words refer to in the ext?	Not to limit the selection to referents that immediately follow their 2 to 4 words
5	OR	<u>Title</u> *Copy the title you think is the most appropriate. Or Give a title to the text.	Not more than three suggested titles
		<u>General idea</u> *Choose the general idea of the text. Or Give the general idea of the text.	Not more than three suggested general ideas
		<u>Type of discourse</u> *The text is... a) descriptive b) narrative c) argumentative d) prescriptive.	To be given if there is only one type of discourse in the text.

2- Text Exploration

Type of activities			Number of activities
1	OR	<u>Lexis</u> *Find in the text words or phrases that are closest in meaning to the following.	2 to 3 words to be found Specify the paragraph for each word.
		* Find in the text words or phrases that are opposite in meaning to	
		*Match the following words with the corresponding synonyms/opposites/definitions.	
		* Find words in the text whose definitions follow.	
2	OR	<u>Morphology</u> *Complete the chart as shown in the example.	3 to 4 words(from the text) to be derived/divided into roots and affixes
		*Divide the following words into roots and affixes.	
		*Give the opposites of the following words keeping the same root.	
3	OR	<u>Grammar</u> *Rewrite sentence B so that it means the same as sentence A .	2 to 3 sentences A passage with 4 verbs in brackets not more than 2 questions
		* Combine each pair of sentences with one of the connectors provided. Make changes where necessary.	
		*Combine each pair of sentences with the connectors given between brackets. Make changes where necessary.	
		*Give the correct form of the verbs in brackets.	
4	OR	*Ask questions which the underlined words answer.	4 words 2 pairs
		<u>Phonology</u> *Classify the following words according to the pronunciation of the final /S/.	
		*Classify the following words according to the pronunciation of the final /ed/.	
		*Classify the following words according to the stressed syllable.	
		*Classify the following words according to the number of the syllables.	
		*Circle or write the silent letter in each of the following words.	
		*Match pairs that rhyme.	

5	OR	<u>Discourse</u>		
		*Fill in the gaps with words from the list given.		4 gaps
		*Fill in the gaps with only 4 words from the list given.		Provide more words than gaps.(6 words)
		*Fill in the gaps so that the text makes sense.		No words given
		*Reorder the following sentences to get a coherent passage.		4 statements
		*Imagine what A or B says, and complete the following dialogue.		3 utterances

3- Written Expression

PART 2 – WRITTEN EXPRESSION	
<u>Both topics should be stated in terms of a situation of communication in which the candidates reinvest and make use of the appropriate resources.</u> <u>Whether guided or free, the statement of the communication situation should comprise the following aspects :</u> • A statement of the problem situation. • The target audience • The objective • The task • The length of the composition <u>Example:</u> <u>Choose one of the following topics:</u> <u>Either Topic 1: (Guided related to the theme of the reading passage)</u> In your city, statistics show that most consumers are unaware of the negative effects of the fake goods they buy. As a member of a consumers' association write a newspaper article of about 80 to 120 words in which you sensitise your fellow citizens about the dangers of fake goods on their health, security, welfare and environment. <u>Make the best use of the following notes : (...)</u>	

Or topic 2: (free and related to one of the other themes in the curriculum)

In your city, statistics show that most consumers are unaware of the negative effects of the fake goods they buy. As a member of a consumers' association write a newspaper article of about 80 to 120 words in which you sensitise your fellow citizens about the dangers of fake goods on their health, security, welfare and environment.

The situation problem	In your city, statistics show that most consumers are unaware of the negative effects of the fake goods they buy
The target audience	fellow citizens
The objective	sensitise fellow citizens / consumers about the dangers of fake goods on their health, security, welfare and environment.
The task	write a newspaper article
The length of the article	about 80 to 100 words

Source: The English Language Test-Design Guide for the BAC Exam (October,2017)

Appendix F

English Language Examination BEM 2014

الجمهورية الجزائرية الديمقراطية الشعبية	وزارة التربية الوطنية
الديوان الوطني للمسابقات والامتحانات	امتحان شهادة التعليم المتوسط
دورة: جوان 2014	اختبار في مادة: اللغة الإنجليزية
المدة: ساعة ونصف	

BRAZIL

Brazil is organizing the Football World Cup which will take place from 12th June to 13th July, 2014. Thirty two countries are participating in this 20th edition. Among them is Algeria, Brazil is hosting this competition for the second time, The first one was in 1950. This country won five times the world cup (1958, 1962, 1970, 1994 and 2002). Do you know Brazil?

It is a big country in the south of America; its area is 8 550 000 square kilometres, ranking 5th in the world. It has got a population of about 190 million inhabitants with a density of 23 persons per square kilometre. Its capital city is Brasilia. Rio de Janeiro and Sao Paolo are the main cities. Portuguese is the official language.

The world's largest tropical rain forest covers most of northern Brazil. The Amazon River goes through this vast forest where a variety of animals like pumas, jaguars, rare bush dogs and foxes live.

Brazil is famous for the coffee production, Samba Music, Rio de Janeiro Carnival and football.

Adapted from: Encarta kids 2008 and www.fifa.com.

PART ONE (14 points)

A/ Reading Comprehension (07 points)

Read the text carefully and do the following activities:

Activity One (02points): Answer the following questions.

- 1- Is Brazil hosting the 2014 Football World Cup?
- 2- Did Brazil win the Football World Cup in 1974?

Activity Two (03 points): Complete the table with information from the text.

Country	Situation	Number of people	Area	Capital	Language	Famous sport
Brazil						

Activity Three (02points): Match the words in column A with their equivalents in column B.

A	B
1- big	a- well-known
2- inhabitants	b- principal
3- main	c- large
4- famous	d- people

B/ Mastery of Language (07 points)

Activity One (02 points): Reorder the following words to get coherent sentences.

- 1- large / a / Algeria / country / is / , /
- 2- Algeria / The fennec / is / animal / , / famous / in / a /

Activity Two (03 points): Write the correct form of the verbs in brackets.

The Algerian team (to participate) in the 2010 Football World Cup, which (to take) place in South Africa. The Algerian people (to support) the team in the 2014 Brazil World Cup.

Activity Three (02 points): Pick out from the text four words and complete the table according to the pronunciation of their final - s.

/ s /	/ z /	/ iz /
1-	1 -	1 -
2-		

PART TWO: Written Expression (06 points)

A Brazilian friend of yours knows that Algeria is going to participate in the 2014 Football World Cup. He / she wants to have more information about Algeria.

Write him / her a letter telling him / her about:

- Situation / area / population
- Capital / main cities
- Famous for

DO NOT SIGN YOUR NAME.

Appendix G

English Language Examination BEM 2019

الجمهورية الجزائرية الديمقراطية الشعبية
الدewan الوطني للامتحانات والمسابقات

وزارة التربية الوطنية
امتحان شهادة التعليم المتوسط
اختبار في مادة: اللغة الانجليزية

2019
المدة: ساعة ونصف

The kind of food we eat depends on which part of the world we live in, or which part of the country we live in. For example, in the South of China they eat rice, but in the North they eat noodles. In Scandinavia, they eat a lot of herrings, and the Portuguese love sardines. But in Central Europe, away from the sea, people don't eat so much fish; they eat more meat and sausages. In Germany and Poland there are hundreds of different kinds of sausages.

In North America, Australia and Europe there are two or more courses to every meal and people eat with knives and forks. In China, there is only one course, all food is together on the table and they eat with chopsticks. In parts of India and the Middle East people use their fingers and bread to pick up the food.

Nowadays, it is possible to transport food easily from one part of the world to another. We can eat what we like at any time of the year. Our bananas come from Caribbean or Africa; our rice comes from India or the USA; our strawberries come from Chile and Spain.

Food is a very big business. But people in poor countries are still hungry and people in rich countries eat too much.

Adapted from "Headway Oxford 2004"
"Food Around the World" - reference: atlas.sk

Part One (14pts)

A. Reading Comprehension (07 pts)

Read the text carefully and do the following activities:

Activity One: Answer the following questions. (03 pts)

- What do people like eating in Portugal?
- How do Chinese people eat?
- Is it difficult to take food from one country to another?

Activity Two: Complete the table with information from the text. (02 pts)

Countries	Food
1. _____	1. _____
2. _____	2. _____
3. _____	3. _____
4. _____	4. _____

Activity Three: Match the words in column A with their opposites in column B. (02 pts)

A	B
1. different	a. whole
2. part	b. die
3. much	c. similar
4. live	d. less

B. Mastery of Language (07pts)

Activity One: Fill in the gaps with the suitable words. (02 pts)

➤ minerals - health - vegetables - low

Fruits and are well-known for being good for Most vegetables are in calories but high in vitamins, and fibres.

Activity Two: Turn the following sentences into the negative. (03 pts)

1. People in the North of China eat noodles.
2. Couscous is served in most Algerian houses on Fridays.
3. In the past, people used to eat healthy food.

Activity Three: Write the following words in the right box according to their final consonant sound. (02 pts)

➤ chopsticks - fingers - depends - parts

/ s /	/ z /
1.	1.
2.	2.

Part Two: Situation of Integration (06 pts)

Your English pen-friend has just sent you an email telling you about the eating habits of his/her country.

Write an email in which you tell him/her about the Algerian eating habits.

Write about:

- ✓ The different kinds of food Algerians eat.
- ✓ Algerian eating manners.
- ✓ Some traditional dishes.

DO NOT SIGN YOUR NAME

Appendix H

English Language Examinations BAC 2014



الجمهورية الجزائرية الديمقراطية الشعبية

الديوان الوطني للامتحانات والمسابقات

دورة: جوان 2014

وزارة التربية الوطنية

امتحان بكالوريا التعليم الثانوي

الشعب: علوم تجريبية، رياضيات، تقني رياضي، تسيير واقتصاد.

المدة: 02 س و 30 د

اختبار في مادة: اللغة الانجليزية

على المترشح أن يختار أحد الموضوعين التاليين
الموضوع الأول

Part One: Reading:

15 points

A. Comprehension

08 points

Read the text carefully then do the following activities.

The study of consumer behaviour examines all aspects of consumer's feelings, thoughts and reasons for making particular decisions in purchasing products or services, and also how consumers use and dispose of products. A consumer's choice or practices may be influenced by family and friends, religious beliefs, cultural attitudes, social expectations, professional standards, advertising appeals, or by any combination of these factors.

The most obvious application for knowledge of consumer behaviour is marketing strategy. For instance, understanding that a large number of consumers are on low carbohydrate diets has led to an increasing number of products that are labelled as "Low Carb". But the study of consumer behaviour also has a lot of repercussions on public policy, social marketing and consumer education.

Marketers may examine consumer behaviour using either primary or secondary research. Primary research is conducted by asking a lot of consumers to answer survey questions, either by mail, Internet, telephone or in person. Mail surveys are useful because they are inexpensive and may ask as many questions as desired. Internet surveys can be cheap to set up but exposure to a great deal of aggressive advertising has made many consumers resistant to this method.

Adapted from the Advertopedia Staff

1. Choose the right answer:

The text is:

a- prescriptive

b- narrative

c- expository

2. Are these statements true or false? Write T or F next to the letter corresponding to the statement.

- The study of consumer behaviour examines all aspects of consumer's feelings.
- A consumer's attitude can be influenced by his family and friends.
- Marketers have to answer a lot of questions in surveys.
- Consumers prefer Internet surveys because they are cheap.

3. Answer the following questions according to the text.

- What are the factors that may influence a consumer's behaviour?
- How can marketers examine consumer behaviour?

4. In which paragraph is it mentioned how to check the consumer's behaviour?

5. Who or what do the underlined words refer to in the text?

a. that (§2)

b. they (§3)



B. Text Exploration

07 points

1. Find in the text words or phrases that are closest in meaning to the following.
 a. comportment (§1) b. evident (§2) c. for example (§2)
2. Complete the following chart as shown in the example.

	Verb	Noun	Adjective
example	to resist	resistance	resistant
	 to consume	behaviour	 educational

3. Ask the questions which the underlined words answer.
 a. Marketers may examine consumer behaviour.
 b. The study of consumer behaviour examines all aspects of consumer's feelings.
4. Classify the words according to the number of their syllables.
 advertising - consumers - obvious - cheap

One syllable	Two syllables	Three syllables	Four syllables

5. Re-order the following sentences to make a coherent passage.
 a. For example, they are entitled to products
 b. Consumers have several basic rights.
 c. They are also entitled to the protection against unsafe food.
 d. whose quality is consistent with their prices.

Part Two: Written Expression

05 points

Choose ONE of the following topics:

Topic One:

In your city you feel that consumers are not protected against the effects of the goods they buy. So, with a group of friends, you decide to create an association of consumers. Write a composition of about 80 to 120 words in which you expose the reasons and objectives of this association.

You may use the following notes:

- **Reasons:** counterfeit / cheap products / lower quality / harmful / not lasting
- **Objectives:** to sensitize the consumers / to protect them / to buy safe products

Topic Two:

Write a composition of about 80 to 120 words on the following:
 Some people think space tourism is just a wild dream and a waste of money. Do you agree with them?



الموضوع الثاني

Part One: Reading:

A. Comprehension

15 points

08 points

Read the text carefully then do the following activities.

Since money came into use in the ancient world, there has been counterfeiting. Early money was in the form of coins with a specific content of gold, silver, or any other metal. Making counterfeit coins was relatively easy. A metal of equal weight could be gold or silver-plated and passed on to the public as genuine.

Counterfeiting paper currency is a highly technical skill that calls for a talented draftsman, photographer, plate maker and printer. Also necessary for the counterfeit money to be convincing are high quality rag paper and the proper kind of ink.

With such technological advances as high-speed cameras and engraving machines, the time needed to produce counterfeit money has been shortened and the quality of the work improved. To combat these problems, the United States government has attempted to make detection of fraud easier by the use of very high quality rag paper and ink. Other nations also use special paper and ink. In much foreign currency a distinctive watermark can be seen by holding the money up to the light.

Counterfeiting is a criminal offence in nearly every country. A convention signed by many nations in Switzerland in 1929, provided for the centralization and cooperation of international police action to combat counterfeiting.

Adapted from Wise Geek .com

1. Circle the letter that corresponds to the right answer.

The text is taken from: a. a magazine b. the Internet c. a newspaper

2. Are these statements true or false? Write T or F next to the letter corresponding to the statement.

- a. Counterfeiting existed since the appearance of money.
- b. Low quality paper and ink are necessary to counterfeit money.
- c. Technological development helped counterfeiters.
- d. Counterfeiting is legal in some countries.

3. Answer the following questions according to the text.

- a. Why was counterfeiting easy in the past?
- b. What has the United States government done to combat counterfeiting money?

4. In which paragraph is it mentioned how to fight counterfeiting money?

5. Copy the title you think is the most appropriate.

- a. The history of Counterfeiting
- b. Money Laundering
- c. Counterfeiting Money

**B. Text Exploration****07 points**

1. Find in the text words or phrases that are opposites to the following.
- a. late (§1) b. more difficult (§3)

2. Divide the following words into root and affixes.

Words	Prefix	Root	Suffix
international			
distinctive			
shortened			

3. Rewrite sentence B so that it means the same as sentence A.
- 1.a. A convention was signed by many nations in Switzerland.
b. Many nations
2.a. "Making counterfeit coins was relatively easy", a historian said.
b. A historian said that.....
4. Classify the following words according to the pronunciation of the final "ed".
passed - shortened - attempted - provided

/t/	/d/	/ɪd/

5. Fill in the gaps with words from the list below.

hide – techniques – financial – and

Money laundering refers to the process of concealing.....(1)..... transactions. Various laundering(2)..... can be employed by individuals, groups, officials(3)..... corporations. The goal of a money laundering operation is usually to(4)..... either the source or the destination of money.

Part Two: Written Expression**05 points**

Choose ONE of the following topics:

Topic One:

Counterfeit is something copied and passed as genuine. What should be done to fight this crime? Use the following notes to write a composition of about 80 to 120 words.

- Sensitize people: mass media / school programmes.
- Organize anti-counterfeiting associations.
- Government: pass legislation / protect consumers.

Topic Two:

Children and teens are often targeted by junk food advertisements.

Write a letter to the director of a fast food company to complain about their effects on those kids.

Sign the letter: John Smith.

على المترشح أن يختار أحد الموضوعين التاليين:

الموضوع الأول

Part One: Reading

(14 points)

A-Comprehension:

(07 points)

Read the text carefully and do the activities.

All over the world, smoking in school is not isolated from what goes on in the streets and in the media. It is affected by advertising, smoking in public places and characters smoking on TV. All these are factors which show it as an acceptable part of our world. Children see parents, friends, teachers and many people doing it. So, this perception is carried into the classroom with them.

The first cigarette is easily taken in a moment of pressure from friends, or classmates. Nobody likes to be bullied, ignored or to be the odd one out. Therefore, if you are offered something, it is better to go with the flow than to say no. Though smoking is, on the surface, kept as a secret at school, it is in reality the worst kept secret at all. Clothes and body smell, hideaways are littered with cigarette-ends, toilets bear testimony with burn marks and pupils are late for lessons.

Teachers often don't know how to tackle the problem by themselves. Currently, many of them are stressed and say that time to address issues like smoking is not available. However, an anti-smoking education is an initiative which needs to be acted on by the whole school staff. Some schools have strict regulations which prevent and punish smokers, but never see the light of day. That's why what should be fostered and preached is a clear assumption that smoking is a real threat to pupils' health and schooling.

Jenny Jacobs. Ray. Soc. Health-February 1993 (Adapted)

1-Are these statements true or false? Write T or F next to the letter corresponding to the statement.

- a- Imitation is the major cause that makes pupils smoke.
- b- Smoking can be kept secret.
- c- It has no harm on pupils' schooling.
- d- Anti-smoking regulations must be applied in schools.

2- Put the following sentences in the order they appear in the text.

- a- Friends' pressure is the principal factor of smoking in schools.
- b- Pupils need adults' advice and guidance to avoid smoking.
- c- Smoking cannot be hidden.
- d- Media incites adolescents to smoke.

3-Answer the following questions according to the text.

- a- Why do school boys and girls smoke?
- b- What should be done to limit smoking in schools?

4- Copy the letter that corresponds to the right answer.

The text is a:

- a- letter
- b- magazine article
- c- survey

5-Who or what do the underlined words refer to in the text?

- a- them(\$1) b- it (\$2) c- which (\$3)

6- Give a title to the text.

B-Text Exploration

(07 points)

1-Find in the text words or phrases that are closest in meaning to the following:

- a- separated(\$1) b- different(\$2) c- encouraged(\$3) d- menace(\$3)

2- Give the opposites of the following words keeping the same root.

- legal acceptable effective hopeful

3- Rewrite sentence (b) so that it means the same as sentence (a).

- 1- a- Governments should interfere to ban bad practices in schools.
b- It's high time.....
2- a- If strict measures are not taken, the educational system will deteriorate.
b- Unless
3- a- I have a strong desire that all governments will ban cigarettes production.
b- I wish

4- Ask the questions which the underlined words answer.

- a- The headmaster is determined to ban cigarettes smoking in his school.
b- Some teachers throw cigarette ends on the class floor.

5-Classify the words according to the number of their syllables.

schooling - policy - affected - smoke

one syllable	two syllables	three syllables

6- Re-order the following sentences to make a coherent paragraph.

- a- more and more school boys and school girls smoke
b- Smoking has spread among students at an alarming rate.
c- because they imitate adults and T.V stars and want to prove their freedom.
d- In spite of its dangerous effects on health,

Part two: Written Expression

(06 points)

Choose ONE of the following topics.

Topic One:

Smoking at school is bad and dangerous for teenagers. Write an article of about 120 to 150 words for your school magazine in which you warn your schoolmates against smoking.

The following notes may help you:

- immediate health hazards: coughs, wheeziness, shortness of breath
- poor academic performance, taking more time outside school
- wasting money

Topic Two:

Write a composition of about 120 to 150 words on the following:

Many children around the world are engaged in child labour. State the causes and suggest the possible solutions to eradicate this phenomenon.

الموضوع الثاني

Part One: Reading

(14 points)

A. Comprehension:

(07 points)

Read the text carefully and do the activities.

Sparta was one of the most important cities in Ancient Greece and it was a very different kind of city state. The life of the people of Sparta was a very strict one, similar to the military. The Spartans were proud, fierce and capable warriors. Young boys were taken from their homes at an early age to begin military training. Young girls were forced to maintain a healthy way of life in order to produce healthy children and were sent to school to learn how to fight and to become soldiers, too.

Most Spartan citizens were either Perioeci (citizens who paid taxes, served in the army and were protected by Spartan laws) or Helots (people from lands conquered and ruled by Sparta who had no rights). Spartan citizens were given land which was farmed for them by the Helots. The Helots were treated as slaves and had to give half their crops to their Spartan master. It was a common belief that the Helots were public property. They were seen as the enemy even though they were actually slaves. The Helots sweated in the fields, but their resentment grew. Finally, they rose up and fought their Spartan masters and the fighting continued for many years. But the Spartans eventually gained victory and so became more powerful.

The Spartans became wealthy through trading in luxury goods of gold, silver and other materials. Besides, they produced beautiful things as their wealth increased. Ivory carvings were desired across Greek lands, bronze-work and pottery were exported to Italy.

Adapted from: greece.mrdonn.org

1. Write the letter that corresponds to the right answer a, b or c.

The text is taken from a:

- a. magazine b. website c. newspaper

2. Are these statements true or false? Write T or F next to the letter corresponding to the statement.

- a. Only Spartan male citizens were trained to fight.
b. The Perioeci class was a wealthy one.
c. The Helots uprose against the Spartans.
d. No great works of art came from Sparta.

3. In which paragraph is it mentioned that...

- a. the Helots were considered a permanent threat by the Spartans?
b. the Spartans led a severe military life?

4. Answer the following questions according to the text.

- a. What was the Spartan's life like?
b. How were the Helots treated?
c. What made Sparta a rich city state?

5. Copy the letter that corresponds to the right answer.

The text is: a. prescriptive b. narrative c. argumentative

6. Who or what do the underlined words refer to in the text?

- a. one (§1) b. they (§2) c. their (§3)

B. Text Exploration**(07 points)****1. Find in the text words or phrases that are closest in meaning to the following:**

- a. violent(\$1) b. governed(\$2) c. hatred(\$2) d. commerce(\$3)

2. Complete the chart as shown in the example.

	Verb	Noun	Adjective
example	to differ	difference	different
		strength	
	to create		
			free

3. Combine the pairs of sentences using the connectors provided. Make the necessary changes.

provided that - as well as - although - because.

- a. Young girls were forced to maintain a healthy way of life. They had to produce healthy children.
b. The Helots fought for their freedom. They were defeated.
c. Our economy flourishes. We enhance agriculture.

4. Give the correct form of the verbs between brackets:

- a. If the Spartans hadn't given much importance to their soldiers they (not/to make) a strong army.
b. After the Spartans (to defeat) the Helots, they became much more powerful.

5. Classify the following words according to the pronunciation of the final 's'.

boys - warriors - barracks - businesses

/s/	/z/	/tʒ/

6. Fill in the gaps with words from the list.

Spartans - time - army - obedience - emphasized - girls

Ancient Sparta gave such a big importance to its ...(1)... that all the boys were brought up in a way that ...(2)... their physical fitness, courage and ...(3).... Only very little ...(4)... was devoted to leisure or family life.

Part Two: Written Expression**(06 points)****Choose ONE of the following topics.****Topic One:**

Using the following notes, write a composition of 120 to 150 words to describe the daily life of the people of ancient Greece.

- houses: built, clay-brick, stone, more than one room, a courtyard
- food: varied, fruits, vegetables, fish, rarely eat meat only in religious feasts
- clothes: home-made, decorated to represent their city-states
- entertainment: dance, music

Topic Two:

Counterfeiting and piracy lead to negative effects. Write a composition of 120 to 150 words about the impact of such practices on people's life.

Appendix I

English Language Examinations BAC 2019



الجمهورية الجزائرية الديمقراطية الشعبية
وزارة التربية الوطنية

دورة: 2019

الديوان الوطني للامتحانات والمسابقات
امتحان بكالوريا التعليم الثانوي

الشعبة: علوم تجريبية، رياضيات، تقني رياضي، تسيير واقتصاد

المدة: 02 سا و 30 د

اختبار في مادة: اللغة الانجليزية

على المترشح أن يختار أحد الموضوعين الآتيين:
الموضوع الأول

Part One: Reading

(15 points)

A/ Comprehension

(08 points)

Read the text carefully then do the following activities:

Cheating, deception and other forms of unethical behaviour are widespread in business, sports, schools, and other arenas. While the media focus on extreme cases of cheating, less attention is paid to what researchers call "ordinary unethical behaviour."

Stealing from one's employer, or cheating on exams are the results of ordinary people surrendering to the temptation to cheat when confronted with the opportunity. These behaviours and others are costly for businesses and society.

Studies find that under some circumstances most people cheat. They don't cheat as much as they can get away with; rather they cheat up to the point at which they believe they are good. When facing the opportunity to cheat, people experience a conflict between their desire to maintain a positive self-image and to advance their self-interest crossing ethical boundaries. One way to resolve this is to cheat a little, reinterpreting the unethical behaviour as an honest mistake.

Adapted from: 'ethicalsyste.ms.org'

1. Are the following statements true or false? Write T or F next to the letter corresponding to the statement and correct the false one(s).

- Unethical behaviours are limited to one domain.
- Unethical behaviours harm businesses only.
- People do not cheat all the time.

2. Put the following ideas in the order they appear in the text:

- Cheaters don't feel they are dishonest.
- Cheaters go beyond morals when cheating.
- Ordinary people may have the desire to cheat.

3. Answer the following questions according to the text:

- Do the media today report all forms of cheating? Justify.
- What dilemma do people face when confronting the opportunity to cheat?
- How do cheaters justify their unethical behaviour?

4. WHAT or WHO do the underlined words in the text refer to?

- These behaviours (§2)
- which (§3)

B/ Text Exploration

(07 points)

1. Find in the text words or phrases that are opposite in meaning to the following:

a. more (§1)

b. cheap (§2)

c. abandon (§3)

2. Divide the following words into roots and affixes:

Word	Prefix	Root	Suffix
unethical			
researchers			

3. Combine each pair of sentences with the connector given between brackets. Make changes where necessary.

a- People are reminded of moral standards. They don't cheat. (unless)

b- Some employees may deceive their bosses. They want to earn more money. (in order to)

4. Reorder the following sentences to get a coherent passage:

a- Interestingly, others' exemplary ethical behaviour affects their likelihood to behave honestly.

b- In fact, when they see others like them behaving unethically,

c- they are more likely to cheat themselves.

d- Facing the decision to cheat, people look for information about appropriate behaviour.

Part Two: Written Expression

(05 points)

Choose ONE of the following topics:

Topic One:

We sometimes let ourselves get away with cheating that we would blame in others for. We tend to judge unethical behaviours in others but not in ourselves. Write an article of about 80 to 120 words for your school magazine in which you raise awareness about the topic.

Make the best use of the following notes:

- cheat / unethical / for all
- behave / moral / values
- watch / imitate / good behaviour
- avoid / judgement / others

Topic Two:

People believe that TV ads help consumers make good choices about what to buy. Write a composition of about 80 to 120 words in which you express your agreement or disagreement with this point of view.

انتهى الموضوع الأول

الموضوع الثاني

Part One: Reading.

(15 points)

A. Comprehension.

(08 points)

Read the text carefully then do the following activities:

Nightmare Journeys

The journey to a new peaceful life is protracted and unmapped, with no guarantee of safely arriving on Europe's shores, let alone being welcomed. Over the weekend of 14 February, 2,600 people were rescued in the Mediterranean off the Italian island of Lampedusa, near where 360 had died last October. The crossing is said to be the most dangerous in the world.

Criminal gangs are the agents for the journey: there is no travel itinerary, travel insurance, luxurious departure lounges, café's and friendly cabin crew, just criminal gangs who charge a fortune and will beat and abuse anyone who challenges them.

They walk, these frightened men, women, children, often for miles, often barefoot or in plastic sandals; sleep on the streets or in the bush; travel from country to country. They are unwanted, intimidated and exploited; risking rape, abuse and death; every step perilous, every day pregnant with uncertainty.

The men, women and children making, what are by all standards, nightmare journeys, are not responsible for the poisonous environment that they have been forced to live in. They are innocent people, who are simply trying to find a peaceful place where they can live, prosper and bring up their families. In so doing, they are being exploited and mistreated by criminal traffickers, police and bandits alike.

by: Graham Peebles
(www.opendemocracy.net)

1. Are the following statements true or false? Write 'True' or 'False' next to the letter corresponding to the statement.

- a- Journeys to the European shores are secured and fixed by traffickers.
- b- Migrants' worries and fears disappear on their arrival to the destination.
- c- Being in a weak position, migrants could not protest against some abuses.
- d- Illegal migration gives chance to increase abuse and mistreatment.

2. In which paragraph is it mentioned that...

- a- journeys usually end up in tragedies?
- b- when reaching the European shores, migrants face disillusion and hostility?

3. Answer the following questions according to the text:

- a- How is the journey to the 'new' place?
- b- Why is "everyday pregnant with uncertainty"?
- c- Why do border runners, 'Harragas', take the risk to travel abroad illegally?

4. WHAT or WHO do the underlined words refer to in the text?

- a- them (§2)
- b- they (§3)
- c- their (§4)

B. Text Exploration

(07 points)

1. Find in the text words that are closest in meaning to the following:

- a- saved (§1) = b- mistreat (§2) =
c- frightened (§3) = d- flourish (§4) =

2. Complete the chart as shown in the example:

	Verb	Noun	Adjective
Example	to migrate	migration	migratory, migrant
		abuse	tolerant
	to exploit		

3. Rewrite sentence 'B' so that it means the same as sentence 'A'.

- 1.A. Migrants are being exploited and mistreated by criminal traffickers.
B. Criminal traffickers
- 2.A. Youths are aware of the dangers. They take risks.
B. Despite

4. Imagine what 'B' says, and complete the dialogue below:

A: Ah finally! Where have you been? You're no longer attending classes these days. Anything wrong?

B:

A: Out of your mind? How are you moving there?

B:

A: Nonsense! Are you sure you can do it?

B:

A: Wake up poor little dreamer! Illegal migration is not that paradise, but just 'living' uprooted; worse, facing death daily.

Part Two: Written expression

(05 points)

Choose ONE of the following topics:

Topic One:

Though the end of the ‘harragas’ stories seem to be different, the journeys share the same worries and fears. Imagine you were one of the rescued ‘harragas’. In about 80 to 120 words, write the story of your **nightmare** adventure to sensitize prospective ‘harragas’ of the risks they are to face.

You may use the notes below:

- worries and fears before and during
- traffickers' aggressiveness
- hard journey conditions
- failure to reach shore
- rescue provided by coast guards

Topic Two:

Unethical advertising always seeks to misrepresent the product in some way. You were a victim of such practices. Write a newspaper article of about 80 to 120 words to denounce and claim for the consumers' safety.

انتهى الموضوع الثانى



على المترشح أن يختار أحد الموضوعين الآتيين:

الموضوع الأول

Part One: Reading

(14 points)

A. Comprehension

(07 points)

Read the text carefully then do the following activities:

Ultimately, all parts of society must share the responsibility for containing corruption because all are willing or unwilling participants as each corrupt transaction requires a "buyer" and a "seller." The government is responsible for dealing with civil servants who engage in extortion and bribery, but it is businesses and individuals who offer bribes to civil servants to obtain certain advantages.

An active, involved, and empowered citizenry is indeed essential to any anti-corruption campaign. Economic reformers can only achieve real gains when a society works firmly and assumes a shared responsibility in combatting corrupt practices, be it collective or individual.

For their part, government institutions, politicians, and bureaucrats must provide the political will to address all such unlawful practices. While all those who are part of the problem must be part of the solution, it would be unrealistic and cost-prohibitive to attempt to eliminate corruption completely. The aim, therefore, is to attain a fundamental increase in honesty, efficiency and fairness.

In this regard, educating and involving the public is a key in preventing corruption, and this can take a variety of ways such as awareness campaigns led by the media, nation-wide integrity workshops that discuss and collaborate against corruption, and the introduction of appropriate legislation put into effect by a powerful and integral auditing system.

Prevention: An Effective Tool to Reduce Corruption"
(GLOBAL PROGRAMME AGAINST CORRUPTION CONFERENCES)

1. Write the letter that corresponds to the right answer.

1. Fighting corruption should be the duty of...
a. all partners. b. the government. c. businesses.
2. The government is responsible for dealing with...
a. corrupt civil servants b. individuals who offer bribes c. both of them.
3. Eradicating corruption is...
a. easily achievable b. hard to achieve c. prohibited
4. Preventing corruption depends primarily upon...
a. weak auditing systems b. public mobilization c. inefficient laws

2. In which paragraph is it mentioned that...

- a. actual benefits result from the combination of social and economic efforts?
- b. promoting ethical principles is an essential way to reduce corruption?

3. Answer the following questions according to the text:

- a. Why do some companies and individuals get involved in bribery?
- b. From a political point of view, who is in a position to lead the fight against corruption?
- c. What measures should be undertaken to prevent corruption?

4. WHO or WHAT do the underlined words refer to in the text?

- a. who (§1) b. this (§4)

5. Choose the correct answer. The text is...

- a. Narrative b. Descriptive c. Expository d. Argumentative

B. Text Exploration

(07 points)

1. Find in the text words or phrases that are opposite in meaning to:

- a. drawbacks (§1) ≠ b. legal (§3) ≠

2. Complete the chart as shown in the example.

	Verb	Noun	Adjective
Example	to corrupt	corruption	corrupt
	to assume		
			involved
		legislation	

3. Combine each pair of sentences with the connectors given between brackets. Make changes where necessary.

- a. Corruption is a widespread problem. All parts of society must fight corruption. (such...that)
b. Governments are making efforts to eradicate corruption. It is still prevalent in our society. (despite)

4. Circle or write the silent letter in each of the following words:

honesty – campaign – while – fight

5. Reorder the following sentences to get a coherent passage:

- a. That's why since the dawn of human civilization.
b. Ethics is important to every society.
c. men have sought to keep the human conduct in check to preserve the peace of society.
d. as it plays a critical role in shaping the individuals' behaviours.

Part two: Written Expression

(06 points)

Choose ONE of the following topics:

Topic one:

You have heard that a friend of yours was asked to give a bribe in order to get a job that he/she is really in need for. Due to his/her difficult situation, he/she was tempted by this offer. Write a letter of about 80 to 100 words to convince him/her not to commit such an offence.

The following notes may help you:

Bribery is unethical – act of dishonesty –have a guilty conscience –illegal and punishable –life to be ruined –favours corrupt people – contribution to mediocrity.

N.B. sign as Ali.

Topic two:

You are a freelance reporter interested in the phenomenon of illegal emigration. You decided to embark with illegal emigrants to experience the hard conditions of their crossing of the Mediterranean Sea and to find out the reasons behind this risky adventure. Write a web article of about 80 to 120 words to sensitise future adventurers about the dangers of illegal emigration.

انتهى الموضوع الأول

الموضوع الثاني

Part One: Reading

(14 points)

A. Comprehension

(07 points)

Read the text carefully then do the following activities:

Violence against children is not an isolated occurrence, it is everywhere; in every country, community and social group. It may take different forms: physical, psychological, sexual abuse, exploitation, neglect or mistreatment, verbal violence, bullying, or cyber-bullying. Violence is a problem of public health, a violation of human rights with potentially devastating and costly consequences.

There is incontrovertible evidence of the harm of violence to the physical and mental health and development of children, to their ability to learn and build relationships with other people, and to grow up as fully-fledged adults and parents. School bullying and an unsafe school environment are among the reasons for early dropout, reduced school attendance, and deteriorating performance of pupils; they have significant social and economic implications. Moreover, we know that education is of key importance for the subsequent professional fulfilment of every individual and the prosperity of societies.

Although school violence has become a major challenge, it is not the education system in itself that causes it. It merely reflects the prevailing social norms, making them visible to everyone. At the same time, the education system makes it possible to address school violence, but only in partnership among all stakeholders: the school, the children, the parents, and the entire community.

Adapted from: <https://www.unicef.org/bulgaria/en>

1. *Are the following statements true or false? Write T or F next to the letter corresponding to the statement.*
 - a. Violence is a form of children's right infringement.
 - b. Violence has negative consequences on children's development.
 - c. Education is crucial for the development of society.
 - d. The education system is the only responsible for school violence.
2. *In which paragraph is it mentioned that...*
 - a. violence is widely spread?
 - b. school failure is due to different causes?
3. *Answer the following questions according to the text:*
 - a. What consequences does violence have on children's growth?
 - b. Who should contribute to the fighting of school violence?
 - c. How do you think education can help societies to prosper?
4. *WHO or WHAT do the underlined words refer to in the text?*
 - a- They (§2)
 - b- them (§3)
5. *Copy the title you think is the most appropriate.*
 - a. Learning Conditions in Schools
 - b. School Violence
 - c. School Bullying

B. Text Exploration

(07 points)

1. Find in the text words or phrases that are closest in meaning to the following:

- a) destructive (§1) b) progress (§2) c) main (§3)

2. Divide the following words into roots and affixes:

mistreatment – incontrovertible – attendance – unsafe

Prefix	Root	Suffix

3. Rewrite sentence B so that it means the same as sentence A.

- A. If the education system doesn't address the problem, violence will prevail.
B. Unless.....
- A. Stakeholders had better work together to put an end to violence.
B. It's high time.....
- A. Children are permanently bullied at school, consequently they drop out.
B. Children drop out.....

4. Circle or write the silent letter in each of the following word:

everywhere – psychological – right – know

5. Reorder the following sentences to get a coherent passage:

- They are working together with parents and the civil society.
- Pupils have become more and more violent.
- to create a safe learning environment for children.
- The education authorities are deeply worried about this issue.

Part Two: Written Expression

(06 points)

Choose ONE of the following topics:

Topic One:

The entire community is deeply worried about the growing report on violence in schools. Write an article of about 80 to 120 words for your school magazine in which you suggest measures to prevent violence and make out of schools a safer place.

You may use the notes below:

- provide psychological counselling services;
- build positive interpersonal relationships.
- reiterate the school rules.
- encourage students not to act violently.
- reward good actions.

Topic two:

Anyone may be the victim of different abuses and misbehaviours on the social media. Meanwhile, you believe that good manners are just as important in cyberspace as they are in real life situations. Write a composition of about 80 to 120 words to be published in your school magazine telling how internet users should behave when using the social networks.

انتهى الموضوع الثاني

ملخص

تهدف هذه الدراسة إلى معرفة ما إذا كان هناك فرق في النتائج الدراسية بين المتعلمين الذكور والإناث في المستويين المتوسط والثانوي في الجزائر وبالخصوص امتحانات اللغة الإنجليزية في شهادة التعليم المتوسط والبيكالوريا. كما تهدف أيضا إلى دراسة مدى تجانس هذه الامتحانات مع محتويات الكتب المدرسية من حيث الموضوعات وأنواع الأنشطة والتمارين. كل هذا من شأنه أن يؤدي إلى استنتاجات حول ما إذا كان جنس المتعلم مؤثر رئيسي للأداء أو هنالك عوامل أخرى محتملة تؤثر على أداء المتعلمين بغض النظر عن جنسهم. تم اعتماد تصميم بحث وصفي مقطعي باستخدام أدواتي بحث وطريقتين رئيسيتين. أولاً، تم إجراء تحليل كمي ونوعي قائم على المحتوى لكتابي اللغة الانجليزية "On the Move" و "New Prospects" على أساس المواضيع المقترحة في كل وحدة / ملف تعليمي بالإضافة إلى أنواع النشاطات والتمارين. تم إجراء نفس التحليل على أوراق امتحان اللغة الانجليزية لشهادتي التعليم المتوسط والبيكالوريا (شعب العلوم التجريبية واللغات الأجنبية فقط) لعامي 2014 و 2019. كما تم كذلك إجراء تحليل إحصائي لنتائج الامتحانين بولاية برج بوعريريج لعامي 2014 و 2019. مناقشة نتائج الإحصاء الوصفي لنتائج المتعلمين من الجنسين تمت على أساس مقارن، في حين أن الإحصائيات الاستدلالية التي استخدمت اختبارات من نوع z-test أخذت بعين الاعتبار فقط عينات من الممتحنين الذكور والإناث الذين تم اختيارهم عشوائياً باستخدام برنامج SPSS. أظهرت النتائج أن هناك درجة عالية من التجانس بين محتوى أوراق اختبار اللغة الإنجليزية لشهادتي التعليم المتوسط والبيكالوريا وكذا الكتب التعليمية المستخدمة من قبل المتعلمين الجزائريين من حيث المواضيع وأنواع الأنشطة والتمارين. كما كشف التحليل الإحصائي أنه لا يوجد فرق بين متوسط علامات المتعلمين والمتعلمات في نتائج امتحاني البكالوريا في شعبة اللغات الأجنبية في كلا الدورتين. علاوة على ذلك، على الرغم من أن تحليل نتائج الاختبارات لعينات شهادتي التعليم المتوسط والبيكالوريا لشعبة علوم الطبيعة والحياة بين وجود فارق في النتائج، إلا أن حجم التأثير كان صغيراً جداً (شهادة التعليم المتوسط: $d = 0.07 / 0.1$ ، شعبة العلوم: $d = 0.07$ ، $d = 0.04$ سنتي 2014 و 2019 على التوالي) مما يعني أن حجم الاختلاف ضئيل ويمكن إهماله. تؤدي هذه النتائج إلى اعتبار أن متغير الجنس بمفرده لا يحدد إمكانية ومعدل نجاح الطلاب في الاختبارات والامتحانات. ومن ثم، تمت الإشارة إلى بعض المتغيرات الخارجية المحتملة التي يبدو، وفقاً لنتائج التحليل، أنها عوامل ذات صلة تؤثر على الإنجازات الدراسية للمتعلمين.

Résumé

Cette étude vise à savoir s'il existe une différence significative entre les élèves masculins et féminins dans les performances des collèges et lycées algériens aux examens de langue anglaise du BEM et du BAC. Il vise également à étudier le degré d'homogénéité des épreuves du BEM et du BAC avec le contenu des manuels en termes de thèmes et de types d'activités. Tout cela conduirait à des conclusions quant à savoir si le sexe est un indicateur majeur de performance ou s'il existe d'autres facteurs possibles qui influencent les performances des apprenants quel que soit leur sexe. Une conception de recherche descriptive et transversale a été adoptée en utilisant deux principaux outils et procédures de recherche. Dans un premier temps, une analyse quantitative et qualitative basée sur le contenu des manuels « En mouvement » et « Nouvelles perspectives » a été menée sur la base des sujets traités dans chaque unité/dossier en plus des types d'activités. La même analyse a été appliquée aux examens BEM et BAC (filieres scientifiques et langues étrangères uniquement) pour les années 2014 et 2019. Ensuite, une analyse statistique des candidats BEM et BAC masculins et féminins dans la Wilaya de Bordj Bou Arreridj dans les années 2014 et 2019 a été réalisée. Les statistiques descriptives des deux populations de candidats ont été discutées sur une base comparative ; tandis que les statistiques inférentielles utilisant deux échantillons de tests z ne considéraient que des échantillons de candidats masculins et féminins (des populations de 2014 et de 2019) sélectionnés au hasard à l'aide du logiciel SPSS. Les résultats ont montré qu'il existe un degré élevé d'homogénéité entre les épreuves d'examen de langue anglaise du BEM et du BAC et les manuels d'instruction utilisés par les élèves algériens en termes de sujets et de types d'activités. L'analyse statistique a révélé qu'il n'y a pas de différence entre les scores moyens des apprenants masculins et féminins dans les résultats des candidats au BAC en langues étrangères au cours des deux années. De plus, bien que cela n'ait pas été correct en analysant les résultats des deux échantillons de tests z pour le BEM et les scores des candidats au BAC de la filière scientifique, la taille de l'effet était très faible (BEM : $d = 0,07/0,1$, filière scientifique du BAC : $d = 0,07 / 0,04$ en 2014 et 2019 respectivement), ce qui signifie que l'ampleur de la différence est insignifiante et peut être négligée. Ces résultats conduisent à supposer que la variable sexe prise seule ne détermine pas la possibilité et le taux de réussite des élèves aux tests et examens. Par conséquent, le chercheur a fait référence à d'autres variables externes possibles qui, selon les résultats de l'analyse, semblent être des facteurs pertinents influençant les résultats scolaires.