

---

**| RESEARCH ARTICLE**

## **An Evaluation of the Literature and Scientific Streams National Baccalaureate Exam Main Sessions' Writing Construct: 2014 to 2018 Academic Years**

**Essaid ELGHAZI<sup>1</sup> ✉ and Abdallah GHAICHA, Ph.D<sup>2</sup>**

<sup>1</sup>*EFL Secondary School Teacher & PhD student, Ibn Zohr University, Agadir, Morocco*

<sup>2</sup>*Associate Professor of Applied Linguistics, TEFL, Educational Assessment, Evaluation, and Policy, Ibn Zohr University, Agadir, Morocco*

**Corresponding Author:** Essaid ELGHAZI, **E-mail:** [essaid.elghazi@gmail.com](mailto:essaid.elghazi@gmail.com)

---

**| ABSTRACT**

The dearth of research studies on EFL assessment in Morocco, as well as the proliferating concerns of EFL in-service practitioners about the quality of the National Baccalaureate Exam of English as a whole, contributed to the Evaluation of the Literature and Scientific Streams National Baccalaureate Exam Main Sessions' Writing Construct between 2014 and 2018, as 2014 is the launch date of a novel guiding guidelines of the National Baccalaureate Exam, a phase that encompasses the inception of the Strategic Vision of the educational system 2015-2030. Henceforth, four objectives were set to lead the study: (1) dissect the inconsistencies of the Exam writing construct, (2) investigate the constituents of the usefulness of the construct in light of previously conducted studies and existing theoretical frameworks, (3) postulating novel ideas to better the quality of the concerned construct usefulness and, (4) enlightening the interventionists in the field about it. Hereby, the researcher adopted a mixed-method design using (1) an evaluation grid to carry out document analysis and (2) an online questionnaire to probe deeper into in-service EFL practitioners' opinions. So, 118 participants, teachers, and supervisors fully completed the questionnaire. The Likert-scaled questionnaires and document analysis checklists that were adopted were confined to Microsoft-Excel and SPSS software statistical graphs and charts. In brief, the study resulted in a plethora of recommendations that would serve to renovate and fortify the usefulness of the exam writing construct.

**| KEYWORDS**

Assessment- Evaluation –Testing -Test usefulness- Construct.

**| ARTICLE INFORMATION**

**ACCEPTED:** 01 July 2024

**PUBLISHED:** 25 July 2024

**DOI:** 10.32996/bjtep.2024.3.2.9

---

### **1. Introduction**

In August 2018, in his speech, his majesty the king Mohamed VI tapped on one of the highest priorities of the Moroccan people, which is the productivity and accountability of our educational system and its interventionists. Therefore, his speech has, along with current observation inquiries, re-affirmed the need to investigate the usefulness of the Moroccan National Baccalaureate written exam of English as an Exit Exam upon which decisive judgements are made about the examinees' future (The National Charter of Education, 1999).

In 1983, the Moroccan Association of Teachers of English (MATE) devoted its yearly meeting to "Testing in EFL," targeting several ramifications of the problem by educational pundits, researchers, and experts from Morocco, the United States, and Britain. The core was the betterment of testing and evaluation in EFL classes in Morocco, which had seemed, theoretically, plausible since then. Nevertheless, students' abilities are still tested in almost the same old way regardless of the recurrent reforms or, ironically, 'reforms of the reforms,' which have become a fact, if not a phenomenon, in our educational system (Ghaicha, 2018).

Moreover, proscribing the pedagogical guidelines of EFL teaching (M.E.N, 2007, 2009) seems to trace a safe path toward the greater good of our nation. However, the status quo implementations of the embellished recommendations that have been produced in the ministry corridors are to be questioned with regard to the outcome.

The writing construct of the Moroccan National Baccalaureate Exam of English is one of the building blocks of the exam that needs to be investigated in light of its alignment with the far-reaching goals of the National Charter of Education and the Pedagogical Guidelines of EFL teaching noted above and the theoretical conceptualization of the construct of writing.

The concern of the current study is probed by the widespread complaints of many practicing teachers about the usefulness of this construct and its conformity to the blue prints of the Moroccan English curriculum.

In a nutshell, the evaluation of the writing construct of the Moroccan English National Baccalaureate Exam main session of Scientific and Literature streams between the years 2014 and 2018 is going to be not only the meanwhile study but also a gateway towards the investigation of novel ways to better upgrade the quality of our exam engineering measures in the sense that high quality engineering measures should ensure efficient use of the right technical input needed to achieve the desired educational outcomes; the better engineered an education system is, the more capable it will be in producing relatively better education results (Ghaicha, 2018; Spolsky, 1998; Shohamy, 2001).

### **1.1 Rationale of The Study**

Assessment and testing practices of a given country are double-edged as they can lead the country into a better destination wherein the pursuit of the social good thrives or, unfortunately, to a bitter fact that puts both the future of society and the credibility of its educational system at stake.

The conceptual shift from the traditional views of student assessment to more contemporary ways of assessment is, theoretically, very promising and inspiring as well (M.E.N, 2007, 2014), but it seems to be, practically, taking a very slow motion with regard to the whole building block formats of the Moroccan English national baccalaureate written exam as a criterion-referenced and high-stakes one. This is embellished by classroom and school ethos, which seem to uphold the assessment of learning while largely compromising the practice of assessment as learning. Hence, advocacy for the new set of assessment ideas and the new assessment culture in which assessment is used as a tool for learning appear to be untransformed into practice (Matseko, 2018).

Language tests play a powerful, if not detrimental, role in many people's lives, acting as gateways at important transitional moments in education, employment, and in moving from one country to another (McNamara, 2000; Amrous & Nejmaoui, 2016; Lazrak & Yechouti, 2017). Therefore, the evaluation of the writing construct of the Moroccan National Baccalaureate Exam, main session, between the academic years 2014 and 2018 is going to be a meticulous investigation of both the usefulness of this construct over the aforementioned span of time through an in-depth analysis of every academic year's exam writing construct usefulness components, along with questioning and analyzing the attitudes and opinions of the hands-on practitioners namely teachers and supervisors.

### **1.2 Research Purpose and Objectives**

Though ornate and appealing Moroccan guidelines for EFL teaching in high school may appear, the majority of teachers are still unhappy about the state of EFL evaluation in Morocco (Melouk, 2001; The Higher Council of Education, Training and Scientific Research, 2015, 2019). Hence, this stance may be related not only to the caveats of current practices but also to a concern with the pedagogical value of the contemporary testing and evaluation methods in such a detrimental exam. So, the current investigation seeks to reach the following objectives:

1. Uncovering and demystifying the shortcomings of the Moroccan English Baccalaureate Exam writing construct of both Literature and Scientific streams.
2. Analyzing the components of the usefulness of the Baccalaureate Exam Writing construct in light of well-documented previous studies and widely recognized theoretical frameworks on the one hand and its constructive alignment, which is yet to be attested, with the official recommendations.
3. Postulating novel ideas to elevate the usefulness of the writing construct in the Baccalaureate National Exam of English.

### **1.3. Research Questions**

Methodologists strike a difference between research problems and research questions. The former refers to some types of difficulty that a researcher encounters or experiences during his or her study of a given topic or phenomenon and for which s/he seeks to find a solution (Kothari, 2004). However, the latter refers to some "specific question[s] asked in the course of investigation to which

a specific answer or set of answers is sought...before arriving at possible hypotheses" (Tavakoli, 2012, p. 49). Furthermore, "good research questions need to address interesting issues" (Dornyei, 2007, p.7). So, to evaluate the usefulness of the writing construct of the Moroccan National Baccalaureate Exam of English, this study tends to answer the following questions:

- Does the National Baccalaureate Exam writing construct reflect a solid theoretical background?
- Does the National Baccalaureate Exam writing construct practically comply with the theoretical guidelines and the specifications set forth by the ministry of education and an attested theory of assessment?
- Are practitioners and hands-on professionals satisfied with the writing construct of the Moroccan National Baccalaureate Exam?
- Does the writing construct of the National Baccalaureate Exam test what should be tested fairly and equitably?

#### **1.4. Significance of the Study**

Regardless of the abundance of educational research in our country years ago up to now (Ouakrim 1986; Ennaji, 1987; Melouk, 1996; Bouziane, 2016; Lazrak & Yachouti, 2017; Ghaicha, 2018), research outcomes were most of the time kept to the researchers themselves, shelves of libraries or international research journals sites in the sense that educational authorities in our country do not heed it much attention but merely launch new reforms as the case of the recent strategic vision 2015-2030.

However, this study will be a plus to the research bank of educational research in Morocco as it tried to probe deeper into the usefulness of the writing construct of the Moroccan National Baccalaureate Exam of English as a high-stakes test and unveiled the recondite credibility and professionalism of those in charge of constructing the written version of the exam that is administered to masses of second year Baccalaureate students nationwide.

## **2. Synoptic Literature Review**

A bulk of available studies, theses, books, journals, articles, and, communiqués ...etc, were dissected so as to serve the purpose of constructing a gearing literature review that would help us address the topic of the study.

Accordingly, the literature review will tap into various key issues related to educational assessment research, notably educational assessment terms definitions, a historical glimpse of language testing, principles of language assessment, types of tests, and test development process. Thus, the researcher will resort to sifting through various sources to assess the writing construct, the development process of a writing test, scoring scales for a writing test, and testing of English as a foreign language in Morocco.

### **2.1 Language Assessment Conceptual Synopsis**

The intertwining relationships between the main concepts of educational assessment make it hard for anyone who attempts to strike a clear borderline between them. Thus, most people who are unfamiliar with terms of educational assessment seem to be confused with the terminology (Ghaicha, 2016; Iliya, 2014).

Bachman (as cited in Ghaicha, 2016, p. 213) states that assessment is commonly used with a variety of different meanings. Since its utility, there are occasions wherein students are assessed either formally through systematic and planned techniques or informally through various indirect ways and techniques.

Notably, assessment is broader than measurement and evaluation despite their intertwining nature (Ghaicha, 2016; Mundrake, 2000) because it includes all kinds of ways to sample and observe students' skills, knowledge, and abilities.

In short, assessment is "...any of a variety of procedures used to obtain information about student performance" (Linn & Gronlund, 2000, p.32). So, it takes three time frames: Pre-instruction, while- instruction, and post-instruction assessment (O'Shea, 2009).

Thus, the former is done through observation of learners' behaviour and reactions in the first sessions or encounters to guard against developing expectations that may distort the teacher's perception of a student. Hence, much of pre-instructional assessment is informal observation (Forgarty & Kerns, 2010).

Santrock (2011) acclaims that "a teacher had better be sensitive, positively, to whether a student is shy or outgoing, has a good or weak vocabulary, speaks and listens effectively, is considerate of others or is egocentric, engages in appropriate or inappropriate behaviour, and focus on the student's nonverbal behaviour for cues that might reveal nervousness, boredom, frustration, or a lack of understanding" (p. 550). Therefore, the teacher's first impressions should be treated as hypotheses, not stable facts, so as not to skew the teaching learning process.

Simply put, teachers and instructors approach assessment from various perspectives according to each one's orientation and reason (Ghaicha, 2016; Santrock, 2011; William, 2011; Airasian, 1994; Linn & Gronlund, 1995; Pelleringo, Chudowsky & Glaser, 2001). However, let us rein ourselves to two most frequently coined terms in assessment: formative and summative assessment.

## **2.2 Formative Assessment**

Currently, one of the buzz words in the field of educational assessment is formative, or while instruction assessment, in the sense that it has become the focal point for all educational interventionists as it diagnoses learners' strengths and weaknesses throughout the whole process rather than after it is completed or what has been referred to as assessment for learning rather than an assessment of learning (Santrock, 2011; Andrade & Cizek, 2010; William, 2010).

Accordingly, formative assessment is not something that happens occasionally; "it is integral to the process of making decisions [...] in teaching" (Iliya, 2014, p.113). William (as cited in Iliya, 2014) recognised that practice in a classroom is formative to the extent that evidence about student achievement is elicited, interpreted, and used by teachers, learners, or their peers, to make decisions about the next steps in instruction that are likely to be better, or better founded, than the decisions they would have taken in the absence of the evidence. Moreover, in formative assessment, judgements about progress and decisions about the next steps take into account the circumstances, past learning, and effort of individual students, as well as what they are able to do in relation to the goals of the work at a particular time. Hence, teachers or instructors take into account all details so as to remedy any potential deficiencies in learners' intake by mending and refining the input that they have been entrusting their learners with.

Last but not least, there are some key components of formative assessment through which instructors hone their learners' skills and competencies (Iliya, 2014). Listed below are but a few of the most crucial practices:

- a. Students are engaged in expressing and communicating their understanding(s) and skills through classroom dialogue initiated by open and person-centred questions.
- b. Students understanding the goals of their work and having a grasp of what is good quality work.
- c. Feedback to students that provides advice on how to improve or move forward and avoid making comparisons with other students.
- d. Students being involved in self-assessment so that they take part in identifying what they need to do to improve or move forward.
- e. Dialogue between teachers and students that encourages reflection on their learning.
- f. Teachers using information about the on-going learning to adjust teaching so that all students have an opportunity to learn.

## **2.3 Summative Assessment**

Summative assessment is "assessment after instruction is finished, with the purpose of documenting students' performance" (Santrock, 2011, p. 551). It aims to measure or summarize what a student has grasped, and this typically occurs at the end of a course or unit of instruction. A summation of what a student has learned implies looking back and taking stock of how well that student has accomplished objectives, but it does not necessarily point the way to future progress (Brown, 2004). As an illustration, final exams in a course and general proficiency exams are examples of summative assessment. In Morocco, English syllabi units' global tests, as well as the National Baccalaureate Exam by the end of the second year of the baccalaureate level, would be an example of summative assessment, so to speak.

In a nutshell, there seems to be no border line between the two main interpretive frames of assessment regardless of their dichotomous appearance, as the case with Swain (1984) and Omaggio (1986). Hence, "they should be thought of as being the endpoints of a continuum" (Carr, 2011, p. 12).

## **2.4 Principles of Language Testing**

With reference to the fact that language assessment is one of the most complex matters in the teaching-learning process (Ghaicha, 2016; Naoua, 2016; Santrock, 2011; Woolfolk, 2016; Bachman, 1990), the current section will try to serve the deal of demystifying the vitality of language tests milestones in the sense that a test would not be a useful one in Bachman's term without them (Bachman and Palmer, 1996).

As gleaned from selective readings of articles, journals, and educational reports, many language teachers harbour a deep mistrust, if not a sort of cynicism, of tests and testers as the fruition of a mutual understanding shortage, sometimes, and a lack of theoretical foundations of the main keys of good testing practices in many instances (Hughes, 2012). Thus, this section will investigate six vital

standards of language assessment, namely Validity, Reliability, Practicality, Accountability, Authenticity, Impact, and Washback or Backwash.

#### **2.4.1 Validity**

Validity has become a buzz word in testing literature significantly in the current approaches wherein validity is no longer, in a sense, the property of the test but of the interpretations that are made of test results (Akbari, 2012; Oller, 1979).

Etymologically, a myriad of definitions have been stated by various scholars and researchers of the term validity. Hence, due to its relation to the questions of "how to test?" and "what to test?", it is defined as the extent to which a test measures what it is supposed to measure (Heaton, 1975; Fultcher, 2010); whereas, Brown (2004) defined it as "the extent to which inferences made from assessment results are appropriate, meaningful, and useful in terms of the purpose of the assessment" (p.22). Furthermore, Bachman (1990) mentioned that "in examining validity, the relationship between test performance and other types of performance in other contexts is considered" (p.236).

Progressively, in the current trend, validity is a unitary concept that concerns the quality of a given test as a unit (Bachman, 1990, 2005, 2007, 2013; Gronlund and Waugh, 2009; Kane, 2013; Babaee, 2017).

So, no matter how difficult validation procedures might be, they are an inevitable part of the test construction process due to the various ramifications of a given test, especially if the test is a high-stakes one (Bachman & Palmer 1996). Test developers, whether they are using a traditional or modern perspective, must apply the test validation processes and mechanics so as to attest that a given test is measuring what it is supposed to measure.

#### **2.4.2 Reliability**

Reliability is considered to be one of the most vital assets of a test in the sense that a reliable test produces consistent scores if it is re-administered at different times (Farhady, 2012; Bachman & Palmer, 1996).

Technically put, reliability is the extent to which a test produces consistent scores at different administrations to the same or similar group of examinees. Therefore, it is defined as consistency of measurement. Sometimes, it can be considered as a function of consistencies across different sets of test task characteristics (Bachman & Palmer, 1996). In other words, reliability is a requirement of test scores due to its investigation of the extent to which measurement is free from errors.

#### **2.4.3 Practicality**

With regard to the previous principles, practicality is vital to the usefulness and strength of a given test. Hence, practicality is defined as "the relationship between the resources that will be required in the design, development, and use of the test and the resources that will be available for these activities" (Bachman & Palmer, 1996, p. 36).

Bachman and Palmer (1996) argued that "the required resources are of three types: human resources, material resources, and time" (pp, 36-37). Pedagogically instigated, a practical test is one whose design, development, and use do not require more resources than are available. In other words, it should be "teacher-friendly".

A teacher should be able to develop, administer, and mark the test within the available time and with available resources (Coombe & Hubley, 2007).

In a nutshell, it should be mentioned that, sometimes, there are trade-offs between ease of development, administering, and scoring a test. For instance, in spite of its easiness of scoring, designing a multiple-choice test is relatively time and energy consuming.

#### **2.4.4 Accountability**

One of the concepts that have always been the centre of attention of many educators in the sixties and seventies is accountability. It started to spread out in the United States in the late sixties, and then it moved to Britain and other European countries at the beginning of the seventies. Equally important, accountability measures are among the assessment methods that can be used to assess educational performance, provide an accurate accountability indicator of high-quality service, and invest efforts to achieve high output quality, procedures effectiveness, and performance improvement.

Tens of studies that have been conducted in the field of education in the Arab world and elsewhere revolve around the utility and necessity of a firm accountability and governance system that will improve and better the fruition of educational practices (Kentab, 2016; Karami, 2013; Almajal, 2009; Alkharaif, 2008; Alharthi, 2008; McNamara & Roever, 2006; Revees, 2004; McAdams, 2003; The Education Commission of the State, 1999; Guth et al, 1999). On the aggregate, accountability in education must be geared towards

holding every interventionist in the field accountable, not just teachers and students but stakeholders as well (Strategic Vision of Education and Training in Morocco 2015-2030).

#### **2.4.5 Authenticity**

Since the current investigation pivots around the usefulness of one of the constructs of a high-stakes exam, it is deemed supportive to go over the concept of authenticity.

Authenticity is a far reaching concept in the domain of language testing due to its concern with the correspondence between the characteristics of a language test task and the features of a target language use task (TLU) (Bachman & Palmer, 1996, p.23).

Similarly important, authenticity provides a tool to investigate the extent to which score interpretations generalize beyond the actual performance of examinees in a language test to the TLU domain or any other similar usage contexts. In other words, Bachman and Palmer (1996) refer to this situation as the degree of match between test task characteristics and an authentic, real-life task. Therefore, Bachman and Palmer (1996) suggested two main questions that should be asked to attest to the authenticity of a given test. Though their general nature, the two questions could be treated as a framework for one's investigation;

- a. To what extent does the description of tasks in the real-life target language domain include information about the setting, input, expected response, and relationship between input and response?
- b. To what extent do the characteristics of the test task correspond to those in a real-life domain?

Finally, Phye (as cited in Ghaicha, 2016, p. 217) contends that "various factors have converged in language testing, among which authenticity is a key factor, and emerged to make classroom assessment one of the most hotly debated issues in education."

#### **2.5 Testing English As A Foreign Language in Morocco**

In their investigation of EFL writing status in Morocco, Tawil, Cerbelle, and Alama (2010) and Amrous and Nejmaoui (2016) as well as Ahmed (2018) affirmed that the current state of both teaching and assessing EFL writing in Morocco is critically at stake especially critical thinking skills manifestation in learners' pieces of writing on the one hand and the deficiency of teachers' assessment literacy on the other (Melouk, 2001). Therefore, assessing EFL writing in Morocco still suffers from various obstacles that impede its usefulness measures and alignment with international standards.

Evaluation measures are to be present in any educational enterprise either directly or indirectly (Melouk, 2001) so as to serve different pedagogical, social, political, and economic purposes. So, it deserves a more vital position in the educational system than has generally been acknowledged.

With regard to the EFL evaluation state in Morocco, diagnostic research was carried out in the academic year 1996 -1997 which was carried out under the auspices of the British Council in Rabat and the Ministry of Education and Vocational Training through the use of questionnaires that targeted the members of the EFL testing units in all the academies and a sample of 563 teachers representing various regions in Morocco. The data that was gleaned from the investigation led to the fact that there are major problems in the relationship between textbooks and the proscribed exams, especially in the first and second year baccalaureate on the one hand and the scarcity of teachers' training in assessment principles it pre- or in-service one (Melouk, 2001).

In the same regard, Bouziane (as cited in Ahmed, 2018) surveyed the assessment of EFL reading and writing in Morocco and concluded that little attention has been devoted to the establishment of a systematic frame of EFL writing assessment, which means that concerted actions must be put forward to better the state of assessing EFL writing of the Moroccan students.

##### **2.5.1 Testing in the National Chart of Education**

Terminologically embellished, the National Chart of Education in Morocco seems to give special credits to the processes of testing and evaluation.

The fifth pillar of the National Chart of Education is ascribed to evaluation and testing, as there are seven articles that tackle evaluation and testing measures 92 – 98 (National Chart of Education, 1999). Notably significant, more weight was given to summative and standardized tests, especially in the first and second year baccalaureate, due to the sensitivity of this phase and its ramifications on society and individuals as well (Articles 96 – 97 – 98 and 103, the National Chart of Education, 1999). Moreover, a national agency was established to account for the quality of evaluation and testing practices and ensure equity and fairness between all (Article 103, the National Chart of Education, 1999). However, the discrepancy between theory and practice is apparently obvious, as it was proved by various sources and research regardless of the recurrent reforms, which sometimes are

sarcastically called “reforms of reforms” (Amrouss & Nejmaoui, 2016; Bedmar, 2014; Tawil et al., 2010; Melouk, 2001). Thus, concerted efforts are to be made in order to secure a firm incarnation of the theoretical tenets of the National Chart of Education in Morocco in the field of education in general and in teaching EFL in particular. As a matter of illustration, article 7 of the National Chart of Education is focalized on striking a balance between all citizens in terms of the quality of teaching input that students must receive in inter-related educational levels as well as insuring the equity of the educational resources be it human, financial, or logistic ones, but the plight here is the fluctual nature of the educational system and the constant state of the theoretical grounds which are regularly subjects to partial refinements that are carried out due to multilateral motives within and outside the country. Similarly, regarding the utility of testing and evaluation in our educational system, article 97 of the chart is concerned with two main points in our system, namely, orienting students towards higher education institutes and universities and then evaluating and nationally ranking high schools on the basis of the annual report that is prepared by the National Agency of Evaluation and Orientation. Nevertheless, the problem here is the evaluation of high schools and the quality of teaching and learning within them, which is, most of the time, based on fuzzy principles that are not clear to the interventionists in the field, let alone to the public. Moreover, if we were to hold accountable high schools or their staff for the failure or success of students on the basis of the National Exams’ grades, we would then be scapegoating teachers and field practitioners to safeguard those backbenchers who just dictate pedagogical texts without an in-depth vision of the real world.

Finally, and perhaps the most noticeable, the National Chart of Education, article 103, set up the foundations for a National Agency of Evaluation and Orientation which serves the developmental research in the domain of human, social, technical and linguistic sciences that are related to educational testing and evaluation, the supervision of educational orientation counselors and coach them alike, the designation of testing and evaluation measures and procedures as well as the establishment of a bank of transparent and reliable test and evaluation samples relaying on the determined official guidelines of teaching in Morocco, the preparation and supervision of national exams that are pre-defined in article 96 of the Chart, ensuring the unification of examinations on the regional scale, the determination of the manner of participation in the international organizations of evaluation, and finally, the preparation of an annual report of its activities and the ranking of schools then release it not only to the concerned agents but to the public as well.

Last, but not least, it is crystal apparent that the National Chart of Educational has established well-grounded foundations for testing and evaluation in Morocco since its inception, but reality has bravely attested the opposite of the embellished texts in a variety of ways among which the Writing Construct of the National Baccalaureate Exam is just the tip of the iceberg that this study will endeavor to uncover its totality.

### ***2.5.2 Testing in the ministerial Guidelines of Teaching English in high school (2007)***

The Guidelines document of Teaching English in Moroccan High Schools appears to be Well-structured and ornate as it taps into some of the main issues and tenets of current practices in teaching, especially the decisive utility of assessment, be it summative or formative. Moreover, more focus was put on the alignment matter between teaching and testing. As an illustration, teaching must lead to testing, not the reverse (Guidelines of Teaching English, 2007).

Accordingly, performance-based assessment is at the heart of the new vision of testing in Morocco, mainly with the adoption of the standard-based approach framework, which is based on five standards. So, assessment in the Guidelines for Teaching English is built on the assumption that language teaching must be related to its social milieu, and so must testing. Simply put, holistic scoring is adopted in order to give more weight to fluency in writing and the communicative side of certain genres of writing (Guidelines of Teaching English, 2007).

### ***2.5.3 Testing in the guiding circulars of the National Baccalaureate Exam of English(2007/ 2010/2014)***

Strikingly put forth, testing in the three circulars is theoretically based on well-established rubrics and scoring methods that technically differ from one section or construct to another, relying on the micro-specificities of each (The National Baccalaureate Exam circular, 2007). Additionally, a great deal of attention was given to what to test, how to score each construct, and how to evaluate the entire performance of the testee.

Of course, there are some novelties from a circular to another, especially the circular of 2014, in order to suit the rapid upheavals in the domain of teaching and learning as there was the addition of some rubrics in writing, especially the ascription of a free writing task (The National Baccalaureate Exam Circular, 2014).

Finally, the testing guidelines of EFL in Morocco are theoretically well-grounded and seemingly more appealing. However, the daily practices may reveal the opposite, as there should be a fair share of teachers’ participation in the testing and evaluation procedures. Teachers must not be only mere recipients of dictated texts that were theorized and embellished in the corridors of the ministry

of the National Education and Vocational Training but rather generators of practically gearing ideas in the educational field because they are the first and utmost direct agents of stability and/or change (Melouk, 2001).

As a consequence, a global revision of testing, evaluation, and assessment practices must be done in order to yield a sound congruence between the written texts that theorize the actual teaching and testing practices in Morocco (Melouk, 2001). In addition, much utility ought to be given to students' critical thinking skills that the ministerial circulars are obsessively infatuated with lately, a tendency which most of the time void of tangible reality once the teaching practices and students' performance are concerned (Amrous & Nejmaoui, 2017). To illustrate this point, teachers are urged to adopt a learner-centered approach to teaching, but, in reality, it is hard for them to implement it in over-crowded classes, low-levels of students, bulky-syllabi that is composed of ten thematic units for all streams, mechanical testing practices which are characterized by 3 to 5 quizzes per semester ...etc. So, the discrepancies are too self-evident to those who are practicing in the field with regard to their daily practices and worries, unlike the educational stakeholders, who fret about no more than embellishing pedagogical circulars and beautifying their yearly reports.

Conspicuously, the shifting tides and inconsistent states of the Moroccan educational system has been the core of various worries of field interventionists. Therefore, the course engine of this study is fueled by both the worries of interventionists and the obvious incongruence between the theoretical grounds, circulars in this context, and the realization of those theoretical guidelines in the National Baccalaureate Exam writing construct as a major part of a high-stakes exam. Hence, the discrepancy between theory and practice with regard to the concerned construct will be the focal issue of the whole paper.

On the aggregate, the general deduction that can be drawn from the entire literature review is that there are such issues that need to be handled firmly and fairly so as to yield accurate results, scores, and interpretations, which would positively impact the lives of testees and society.

Accordingly, the twenty five studies, as well as books that have been consulted and referred to throughout the literature review, reveal that assessment and evaluation are often done furtherly far from the specifications and theoretical grounds that should, presumably, lead the testing process, especially in the third world contexts, Asia, North Africa and Arab world countries (Ahmed & Abouabdelkader, 2018). Moreover, as gleaned from the aforementioned facts, the main obstacles that hinder the usefulness of language tests are multifaceted, but few are to be mentioned here. First, there is a lack of, and in some instances, a lack of, well-grounded assessment literacy on the part of test designers, testing authorities, and teachers. Second, the false alignment between the blue prints of a language test and its actual realization. Finally, the scarcity of research in the field of language testing and assessment. Albeit, in case any corrective and directive research exists, it remains archived on the shelves without implementing its results and recommendations to better the educational status quo in Morocco. Therefore, all the pre-sieved literature sources show that an evaluating study of a primordial standardized test, namely the National Baccalaureate Exit Exam writing construct, is theoretically demanded so that the research bank in Morocco could be consolidated by such endeavor.

### **3. Methodology**

This study gauged the degree of the Usefulness of the Moroccan National Baccalaureate Exam Writing Construct on the basis of sifting the exam papers writing constructs between the 2014 and 2018 main sessions, comparing the educational and pedagogical theoretical frameworks that orchestrate the path of education in Morocco notably the National Charter of Education, the Ministerial Guidelines of EFL teaching in High School, the Strategic Vision of Education Reform 2015-2030 and the Ministerial Circulars with the actual Baccalaureate Exam papers so as to detect the pitfalls and shortcomings of it and suggest potential ways and solutions to remedy the situation.

To address the previously raised inquiries, the adoption of a mixed method research design is more advanced due to the researcher's inclination to gain more freedom to explore various ways to amass accurate data and, therefore, warranted results. Furthermore, as it was affirmed by Dornyei (2007), "...combining both qualitative and quantitative design will back up the potential that the strengths of one method can be utilized to overcome the weaknesses of another method used in the study" (p.46). In short, the undergone approach was opted for to seek convergence and corroboration as well as eliminate bias while cementing the credibility and validity of the findings. Moreover, the security from subjective inconsistencies that this approach grants the researcher strengthens the motive for its adoption by the researcher.

#### **3.1 Context of The Study**

There has been an incessant recurrence of educational reforms in Morocco since the inception of the National Chart of Education in 1999, hitherto in theoretical congruence with the requirements of the World Bank of Development Reports so as to alleviate the actual chasm between the theoretical fantasy and daily reality in the field of education. Nonetheless, the *status quo* of the Moroccan



educational system is bleak, a fact that was warned against years ago (Ghaicha, 2018; Melouk, 2001). Remarkably, Moroccan stakeholders have been madly infatuated with the concept of "Fair governance" and "Accountability" in all public sectors of life, including the educational one. Nevertheless, in Morocco, it is not just a matter of fair or unfair governance and just or unjust accountability, but it is a crisis of content that the educational system is strained with (Tawil, Cerbelle, & Alama, 2010).

In the aggregate, the present study will focus on the analysis and evaluation of the writing construct of the National Baccalaureate Main Session Exam of English in Morocco between 2014 and 2018 so as to gain an in-depth vision of its usefulness.

### **3.2 Research Design**

This study adopted a mixed method research design in order to probe deeper into the core of the investigated topic. As Sugiyono (2010) defines the research method as the scientific way to get the data with certain goals and specific functions, a mixed research design is the most typical way of doing scientifically based research.

Accordingly, an evaluation grid, which is part of document analysis that is used in qualitative designs, was used to account for content and face components of the usefulness of the concerned construct through analyzing both Literature and Scientific streams' exam papers over the defined period of time that starts from 2014 to 2018 so as to get reliable data and therefore valid results. Then, an online questionnaire was used to elicit in-service EFL teachers' and ELT supervisors' opinions about the studied construct.

### **3.3 Participants and Sampling**

Given the typology of the adopted research design, the sampling method combined two divergent ways, probability, and non-probability sampling, in the sense that the targeted population was confined to the convergent usage of the two. Thus, for the vitality and sensitivity of this topic, the number of participants was directed to yielding a large bank of data so that the results, discussions, and interpretations of findings would be sophisticatedly fortified. In such studies, the number of participants should, preferably, exceed the least number of 30 respondents, as it was argued by Dornyei (2003):

From a purely statistical point of view, a basic requirement is that the sample should have a normal distribution, and a rule of thumb to achieve this, offered by Hatch and Lazaraton (1991), is that the sample should include 30 or more people. (p. 74)

Thus, the vitality of the topic necessitated getting in touch with the following groups of hands-on practitioners to get in-depth and insightful data about the issue and, therefore, accurate results that would serve the educational as well as the social good of the Moroccan EFL students. Moreover, through analyzing the reached data from the apparently key participants in the Moroccan educational system, a succinctly scrutinized visionary conclusion about the engineering status of the system was at hand in accordance with what has already been warranted and warned against by the World Bank of Development Report in 2007 about the unstable *status quo* of various MENA countries' educational systems among which Morocco is not an exception. Hence, the engineering state of the Moroccan educational system is at stake in the absence of real basic solutions that would rejuvenate the old practices and mentalities as well as regenerate the engineering intra-treatments from within (WBDR, 2007). In a nutshell, data were gathered with the co-operation of the following interventionists:

1. Supervisors of ELT in Morocco (28 supervisors from divers Local Delegations and Regional Academies of Education).
2. Public high school In-service EFL teachers (90 EFL teachers from all-over Morocco).

Accordingly vital, its compatibility with and appropriateness for reaching the main objectives of this investigation, snowball sampling was adopted in order to gather reliable and valid data. The choice of this technique emanates from the researcher's professional and pedagogical relationships with the net-work of the target respondents on the one hand and its appealing advantages on the other. Snowball sampling is based on the idea that a core group of participants is initially sampled for the research project. The participants are then asked to identify others who might be eligible to participate. Then, this second generation of participants is, therefore, contacted. Consequently, these people, in turn, identify other participants. Hence, the sample, like a rolling snowball, begins to build on itself and increases in size (Vanderstoep & Johnston, 2009).

In short, purposeful sampling was the right option to obtain authentic and valid data.

### **3.4 Instrument and Procedure**

The type of data sought for investigation originates from two main broad sources:

### **3.4.1 Evaluation Grid**

Document analysis, which includes evaluation grids, is a systematic procedure for reviewing or evaluating documents, both printed and electronic material. Corbin and Strauss (2008) affirmed that along with other analytical techniques in mixed method research, document analysis requires that data be examined and interpreted in order to elicit meaning, gain understanding, and develop empirical knowledge. Hence, this study used two exploratory evaluation grids that were composed of thirteen areas of usefulness to check out the degree of absence and presence of the componential elements of the usefulness of the National Baccalaureate Exam writing construct, main session, of both Literature and Scientific streams along with tapping on a variety of official documents to come out with a thorough but precise body of data. The following are the consulted documents that were qualitatively analysed using an evaluation grid:

- a. Baccalaureate Exam paper writing construct 2014 - 2018 (Literature and Scientific Streams; Main Session).
- b. National Charter of Education, 1999, p.35.
- c. Guidelines of EFL Teaching in Moroccan High School, 2007, p.26.
- d. Ministerial Circulars of the National Baccalaureate Exam of English 2010, 2014.

### **3.4.2 Questionnaire**

The researcher opted for a questionnaire as another main instrument to foster the validity of findings that were sifted out from analysing diverse documents on the one hand and the compatibility of such instrument with the objectives of the study, among which reaching out to different EFL pedagogical supervisors and teachers from all educational academies in Morocco was one of the tenets of the meanwhile investigation, on the other. So, as a teacher of English in a public high school in Taroudant Delegation and an active member in National Virtual groups of the Moroccan Association of Teachers of English (MATE), a Google form questionnaire was the most appropriate instrument to gather data from scattered respondents all over the country.

The questionnaire relied on a range of closed-ended questions and only one open-ended question so that the respondents would not have to spend much time with the questionnaire on the one hand and ease the task of data processing and analysing on the other. Nonetheless, the researcher's intent was to use a multi-scale questionnaire to do a good job of the psychometric reliability of findings, but with regard to the workload and time schedules of respondents, who are in-service interventionists, the researcher resorted to closed-ended questions to save the respondents' time and elicit relevant data to such a nationally crucial and sensitive, in some respondents' terms, a topic like the National Baccalaureate Exit Exam, Main session.

### **3.4.3 Validity of Instruments and Findings**

Validating the accuracy of results themselves is much more vital in academic research, regardless of whether it is qualitative or quantitative research (Creswell, 2012).

The validity of findings in this investigation is the counterpart of trustworthiness in qualitative research. So, "validity of findings is manifested if the research instruments actually measure what they were claimed or intended to measure accurately and precisely" (Vanderstoep & Johnston, 2009, p. 27).

Similarly important, the piloting process and item analysis of the questionnaire, which are rudimentary prerequisites to ascertain the practicality and the predictive validity of any useful questionnaire, were carried out with much zeal and cautious guardianship against any potential inconsistencies or pitfalls.

Accordingly, the evaluation grids that were used to evaluate the usefulness of the concerned construct were neatly revised and cross-checked again with the help of two colleagues who are EFL teachers at Aoulouz high school.

In short, the researcher restrained himself from abiding by the ethical standards of not only ensuring the validity of the findings but safeguarding the respondents' secrecy as well in the sense that the researcher reassured the respondents that the information would only and exclusively be used for research purposes. Moreover, respondents' gender, marital status, and other private matters were kept away so that any potential bias against anyone would have neither room in the investigation nor a chance of interference in the results' discussion and interpretation.

### **3.4.4 Piloting of the Questionnaire**

Fundamentally, the piloting process of a given questionnaire or survey denotes the small-scale of trialling questionnaires or surveys that researchers administer to a representative sample of the target population before the main investigation is conducted (Blaxter, Hughes & Tight, 2006). In addition, trial runs of questionnaires allow the researcher to collect feedback about how the instrument works and whether it performs the job it has been designed for (Dornyei, 2002; Selvamuthu & Dipayan, 2018). Therefore, an online

pre-test or pilot run of the questionnaire was initiated in four high schools by emailing the questionnaire to seven teachers synchronously. The concerned high schools are the following: Aoulouz High School in Aoulouz city, Dakhla High School in Ouled Berhil city, Mohammed VI High School in Taliouine district, and Elfadila High School in the Tafingult district. The piloting made sure that the participants did not encounter any difficulty in responding to the questions.

Emphatically vital, the aforementioned high schools were just trialing fields to test the well-formedness and validity of the questionnaire. Thus, teachers who took part in the piloting run were not targeted as main subjects in the main empirical dissemination of the questionnaire.

It should be noted that the researcher was committed to ensuring an equal representation of both genders. Piloting, then, allowed him to check the validity and the practicability of the questions and the time taken to fill in the questionnaire.

As a result, the pilot questionnaire was used to improve and fine-tune the final version and anticipate the response rate to questions respectively.

### **3.4.5 Data Analysis**

Data analysis made use of theoretical backgrounds theorized both by qualitative and quantitative research methodologists as well as the adoption of Microsoft Excel and SPSS software in their simplistic versions, taking into consideration the researcher's fear of any potential statistical inconsistencies that might emanate from the modesty of one's knowledge with regard to advanced options of such software, as well as the utilization of statistics that google questionnaire platform provides the researcher with to numerically analyze and represent quantitative data. A research logbook was ascribed to note-taking while visually analyzing the questionnaire. Then, tabulation, graph forms, and formats were relied on so as to quantitatively sort out data from the Google Forms questionnaire.

To sum up, the aforementioned methodology components were treated carefully to delineate the coming sections in a purposeful manner due to the researcher's rooted respect for the ethics of academic research.

## **4. Results**

Findings and results are couched and presented qualitatively and quantitatively according to their nature so that the presence of validity and reliability standards would be maximized (Brannen, 2016; Vanderstoep & Johnston, 2009).

For organizational reasons, the results and findings shown here are divided into two broad divisions: the first is document analysis of the National Baccalaureate Exam Writing Construct -Main session- between the academic years 2014 and 2018 hand in hand with the related ministerial guiding circulars via utilizing an adopted evaluation grid which is usually used to check a given language test usefulness; the evaluation grid was suggested by Bachman and Palmer (1996). So, since its qualitative-quantitative nature, the grid or checklist is to be attached here in this section for the purpose of easing the reader's task to visualize its convenience utility in this study on the one hand and also fostering the plausibility of the resulting data on the other.

The second division of data is gathered through the use of an online Google Forms questionnaire that is comprised of sixteen items, among which the majority are scaled ones. The questionnaire was concerned with the stance of in-service field interventionists, 90 EFL teachers, and 28 ELT pedagogical supervisors regarding the componential elements of the National Baccalaureate Exam Writing Construct Usefulness. The convergence of the above division was manifested by including the four objectives of this study under the label of evaluating the usefulness of the National Baccalaureate Exam Writing construct – Main Session. Furthermore, to probe deeper into the main strings of the theme, a qualitative consultation of educational documents that were pre-stated and a purposeful questioning of in-service practitioners in the field of education in public high schools in Morocco were carried out to yield unequivocal facts that should support the study's aims by the end.

### **4.1 Analysing the Usefulness of the Baccalaureate Exam Writing**

#### **4.1.1 Construct Through EN Evaluation Grid**

The evaluation grid, which was used by the researcher, is a checklist of usefulness suggested by Bachman and Palmer (1996) and adapted by the researcher to cope with the requirements of the topic under investigation.

#### **4.1.2 Authenticity and Target Language Use Domain Presence and Absence in The Writing Construct**

As gleaned from the evaluation grid, the majority of the writing construct tasks respected the principle of authenticity and TLU domain throughout the investigated years, as 4/5 respected it.

#### **4.1.3 Topic familiarity with regard to the syllabi**

Concerning this principle, the writing construct tasks for all the consulted years completely fulfilled it as we find that 5/5 is the ratio of familiarity presence with regard to the syllabi themes, but interest is another issue that would be checked via the questionnaire.

#### **4.1.4 Prompts and Stimuli That Were Used to Activate Testees' Knowledge**

For this point, all kinds of prompts were absent apart from the recurrent usage of statements, which were used four times throughout the five years, and graphs/images, which were used twice.

#### **4.1.5 Number of The Assigned Writing Tasks**

Significantly clear, one task in the writing construct was used in 2014, but in the remaining years, two tasks were used per year.

#### **4.1.6 Type of Tasks**

When it comes to types of tasks, a balance existed between guided and free tasks in terms of whether students were being given any instructions that would lead them or not. For free tasks, only two cases out of the investigated samples were free ones.

#### **4.1.7 Writing Genres**

An obvious remark is that all genres were neglected or excluded apart from articles and reports that constituted 1/5, respectively, then the frequent usage of e-mails, mainly informal ones, and short paragraphs that represented 2/5 and 4/5 consecutively.

#### **4.1.8 Types and Relevance of The Writing Tasks**

The dominant writing types over the five year period are expository 4/5 and argumentative 2/5, whereas narrative and descriptive represent no more than 1/5 each in the absence of persuasive writing tasks. However, all the tasks proved to be relevant to the scientific stream.

#### **4.1.9 Task(S) Reflection of Construct Definition**

The tasks seemed to partially reflect the definition of the writing construct as it was defined in this paper and in the Guidelines of EFL Teaching in Morocco.

#### **4.1.10 Target Language Use Domain Reference and Practicality Of Tasks**

TLU domain is well represented in most tasks 4/5 except for one, as the tasks were all practical 5/5 with regard to their manifestation in the real world.

#### **4.1.11 High-Order Thinking Skills Presence**

Significantly clear, all tasks necessitated students' synthesis skills 4/5 and, to a little extent, evaluation skills 1/5, but analysis skills were absent throughout the five years.

#### **4.1.12 Standard-Based Approach Standards Presence**

Strikingly appealing, Connection and Comparison were 5/5 present, whereas Community standard was 4/5 present, Communication standard was 3/5 present, but Culture standard was marginalized or overlooked.

Analogically vital, the evaluation grid had uncovered many deficiency points that need to be discussed in the section of discussion along with comparing the findings with facts or conclusions from previously carried out studies in the domain regardless of their scarcity as it was already postulated.

In sum, the evaluation grids made some of the deficiencies of the usefulness of the studied construct crystal clear, mainly practicality, authenticity, and TLU domain, and then impacted learners when it comes to thinking skills.

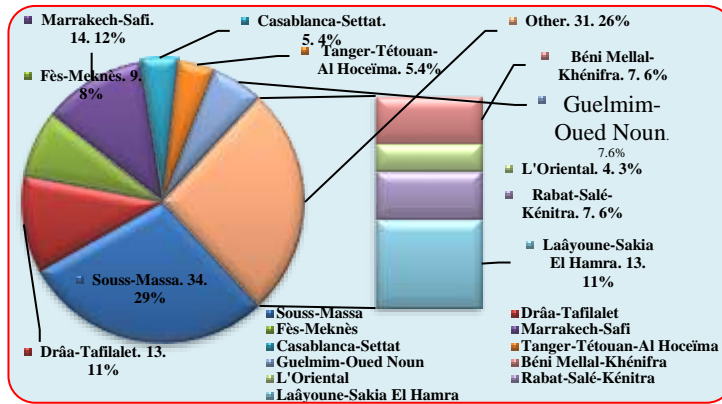
#### **4.2 Moroccan Efl Teachers and Elt Supervisors' Attitudes**

For objectivity, consistency and clarity requirements, the researcher opted for stating findings in statistical states as they were elicited from the respondents using pie charts and a summary table of comments because of their informative nature and plain display of facts. Furtherly crucial, the research design, as well as the researcher's intent to communicate one's findings not only to experts and evaluators of the paper but also to ordinary people, dictated the compartmentalization of findings from discussion (Vanderstoep & Johnston, 2009). As the foregoing exposition should make clear, it would be informatively ethical as it was inspired by Dornyei (2002) to note that demographic, occasionally called private, variables such as gender, age, marital status, and academic achievement were not overlooked or ignored haphazardly but they were avoided so as to maximize the number of respondents

and answerability rates by being straightforward on the one hand and creating a non-threatening milieu to the target respondents on the other.

**4.3 Respondents' regional Academies Belonging**

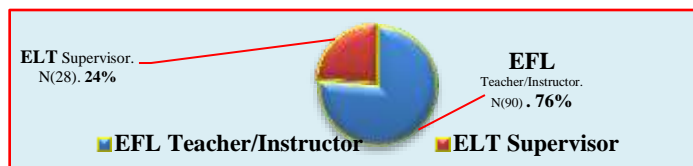
Symbolically significant respondents, teachers, and supervisors who reacted to the questionnaire belonged to various Regional Academies of Education in Morocco, as figure 5 shows. Consequently, the researcher's intent of covering all regions in Morocco was statistically realized.



As gleaned from the above pie chart, the highest portion and percentage of participation was realized in the Souss Massa region, with a total number of 34 respondents constituting 29% of the whole sum of respondents throughout the country. Then, the Marrakech-Safi region came second with a number of 14 respondents and a ratio of 12% of the rest of the respondents; Daraà Tafilalt and Laâyoune-Sakia El Hamra regions both shared respectively equal numbers of 13 respondents with a rate of 11% each, Fès-Meknès region with 9 respondents constituting 8% of the total sum, Béni Mellal-Khénifra, Guelmim-Oued Noun and Rabat-Salé-Kénitra shared again equal numbers of respondents each as the figure above shows 7 EFL practitioners reacted to the questionnaire from each region with a percentage of 6% each. Small numbers of respondents were resulted from the region of Casablanca-Settat and Tanger-Tétouan-Al Hoceïma, with 5 respondents constituting 4% of the total ratio. Finally, L'Oriental showed the least portion of respondents as there were 4 respondents that constituted 3% of the whole. Remarkably, the majority of respondents were from the regions of Souss-Massa and Marrakech.

**4.4 Respondents' Status Or Mission**

EFL practitioners in Morocco, restrictively supervisors, and teachers showed a generous and committed reaction towards the topic of the investigation as 28 ELT supervisors had completely filled out the form with a percentage of 24% of the target population. Similarly important, EFL teachers and instructors were highly engaged in the investigation with a number of 90 teachers, which resulted in 76% of the total range of the participating population.



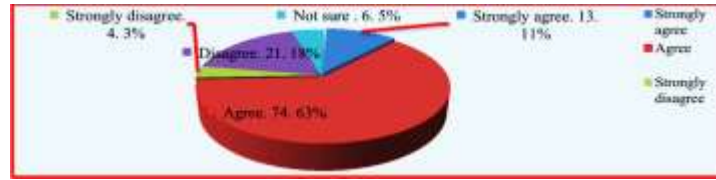
Finally, as it was displayed in the chart above, all 118 participants filled in the questionnaire in a reasonable time, as the majority did the task just after its launching in a time span of no more than ten days.

**4.5 Respondents' Years of In-Service Experience**

Significantly displayed, the 0 to 5 years category was the dominant one with a number of 51 participants that constituted 43% of the total sum. Then, the second category is the 5 to 10 years one, in which 23 participants were enrolled forming 20% of the whole. The third major category is the participants' category of those who had more than 20 year of experience in the field, with a total number of 19 members and a percentage of 16%. So, the category 15 to 20 years came at the fourth rank with 14 members and a ratio of 12%, which was quite an expressive percentage. Finally, the 10 to 15 years category was the smallest range as there were only 11 members, constituting 9% of the total number.

**4.6 The National Baccaureate Exam Writing Construct-Section Tests What Should Be Tested**

The statistical differences between the generated reactions to the above statement were significantly clear. So, as figure 8 below shows, 74 participants agreed with the statement, with a percentage of 63% of the total number. Then, 21 members disagreed with it, constituting 18% of the total sum of the population, followed by the strongly agree category with 13 members, which constituted 11% with regard to the population. Finally, the strongly disagree category with the statement resulted in 4 participants with a percentage of 3% of the whole, while the Not sure category was composed of 6 individuals with a ratio of 5%.



**4.7 The National Baccalaureate Exam Writing Construct-Section Task(S) Reflect(S) Real World Situations (Authentic Tasks)**

The above statement had generated various response rates with regard to the degree of authenticity in Bac Exam Writing construct tasks, as figure 9 below displays. Notably clear, 90 respondents had chosen the "To some extent" value, representing 76% of the population which reveals that the majority has a middle stance position against the authenticity of the Exam.

Then, 17 respondents opted for the judgement value "To a large extent," forming 15% of the total number of respondents.

Statistically appealing, 11 members out of the 118 respondents went for the judgement value "Not at all," making out 9% of the total ratio of responses.

In short, all three categories conveyed logical significance with regard to the general population's number, as every categorical judgement would have a say, then.

**4.8 The National Baccalaureate Exam Writing Construct Task(S) Themes (Topics) Are Interesting To Students With Regards To Their Streams (Literature#Scientific)**

A huge number of respondents went for the value "Somehow interesting," as 78 respondents had chosen it, making a percentage 66% of the total sum of the population, whereas the rest portion was divided between those who opted for the value "Not interesting" who constituted 22 members with a ratio of 19% and then 18 respondents went for the value "Very interesting" forming 15% of the sum total.

In conclusion, the value "Somehow interesting" was significantly large in comparison with the other judgement values, among which the value "Not interesting" was a significant one too.

**4.9 The National Baccalaureate Exam Writing Construct Task(S) Take(S) Into Account Individual Difference (Intelligence, Age, Gender, Etc)**

Regarding the extent to which the National Baccalaureate writing construct task(s) take into account individual differences, respondents had different views, as is warranted in figure 11 below. Numerically expressive, 40 out of 118 respondents went for the value judgement "To some extent," making 34% ratio out of the major total then 38 respondents opted for the value "Not at all," constituting an expressive percentage of 32% regard to the total population and the other judgement values. So, for the judgement value "To a lesser extent," the chart shows that 33 respondents went for it with a ratio of 28% out of the population. Finally, only 7 respondents out of 118 advocated the above statement forming only 6% out of the whole sum.

**4.10 The Design Of The Writing Construct**

Statistically significant, the judgement statement "It needs some refinement" was stated by 65 out of 118 respondents, with a net percentage of 55% regarding the other judgement values. Then, at the second rank, the value judgement "Well-designed" came as it was opted for by 41 respondents out of the rest, making up 35% of the total. Finally, the value judgement "Ill-designed" was ticked by only 12 members out of the total population, which represented a 10% ratio out of the total. In a nutshell, the aforementioned statistics show that a large circle of the population is unhappy or unsatisfied with the actual writing construct task(s) design.

**4.11 The National Bacculaureate Exam Writing Construct Task(S) Have Become Recursive (Repetition Of Specific Genres) Every Year**

A significant number of respondents agreed with the idea of the recursive nature of task types throughout consecutive years. Hence, 61 respondents agreed with the statement, with a percentage of 56% out of the total. In addition, 36 respondents strongly agreed with the statement too, resulting in a 33% ratio out of the target population.

For the ones who disagreed with the statement, 10 respondents disagreed with it, making 9% of the population, whereas only 2 respondents were strongly disagree with the statement, making a range of 2% of the subjects.

Remarkably, all subjects gave their opinions about the above statement, among whom the majority saw variably that repetition became obvious. Therefore, no one abstained from expressing one’s opinion as the “Not sure” value was not used as its ratio was 0%.

**4.12 The National Bacculaureate Exam Writing Construct Task(S) Reflect(S) The Standard-Based Approach (Communication, Community, Connection, Culture, And Comparison)**

This item generated two substantial divisions among the five. Thus, the first striking one is the category of respondents who ticked the “Agree” value, with a total number of 59 respondents constituting a rate of 61% of the total. The second alluring group is the “Disagree,” one with 30 respondents that made up 31% of the subjects. Then, two other groups that respectively stood for the ends of two divergent extremes, namely “Strongly disagree” group with 5 respondents that, made 5% and the “Strongly agree” group with 3 subjects that resulted in a 3% rate out of the sum population respondents.

**4.13 The National Bacculaureate Exam Writing Construct Task(S) Correspond(S) The Theoretical Guidelines Of Tefl In Morocco**

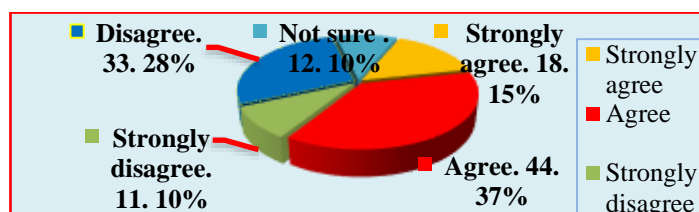
A large number of respondents went for the judgement value “Partial correspondence,” as 75 respondents opted for it, forming the highest portion of 64% of the whole population, whereas 39 subjects opted for the “Full correspondence” judgement value making 33% of the population. Then, only 4 respondents chose the judgement value “Not at all” with a ratio of 3% of the whole sum.



As deduced from the above chart, partial correspondence between the actual hard version of the National Bacculaureate Exam writing construct and the Guidelines for Teaching English as a foreign Language in Morocco was proved by the majority of respondents.

**4.14 The Scoring Scales And Schemes Are Well-Explained To Testees**

With regard to respondents’ reactions towards this item, figure 16 below shows that 44 respondents agreed with the claim that the scoring scales and schemes of the writing construct of the Exam are well-explained. So, they constituted 37% of the total population. However, 33 respondents opted for the “Disagree” value in stark contrast with the first group, representing 28% of the accessed subjects, followed by those respondents who chose the “Strongly agree” value with a total of 18 members that, represented 15% of the population. Then, 11 members strongly disagreed with the claim, making up 10% of the total population. Contrastively, there were 12 respondents who stated that they were not sure about the claim, with a percentage of 10% out of the population.

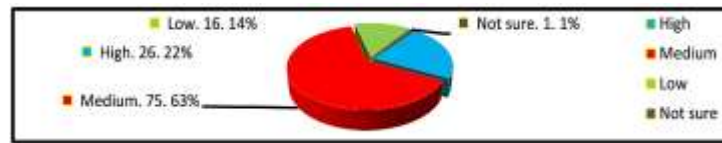


Simply put, the percentage of respondents who proved the claim is relatively high in comparison with the ones who opted for the “Disagree” judgement value.

#### **4.15 Fairness And Quality Of The Exam**

As figure 17 displays, 75 respondents opted for the judgement value "Medium," representing 63% out of the total respondents; 26 respondents chose the value "High," making a percentage of 22% out of the whole sum; 16 respondents chose the value "Low" forming 14% out of the whole participating population and, finally, only one respondent opted for the option "Not sure" making up 1% representation.

Therefore, "Medium" and "High" values got the lion's share concerning fairness and quality of the National Baccalaureate Exam writing construct-Main session.



In conclusion, a positive majority of views was detected about the fairness and quality of the Exam conditions and administration procedures.

#### **4.16 The Extent To Which The Task(S) Of The Writing Construct Target(S) Testees' High-Order Thinking Skills**

More than half of the population had opted for the value judgement "To some extent," as 63 respondents chose it, making up 53% of the main subjects. However, "To the least extent," value judgement got a statistically expressive portion of ticks, too, as 39 respondents went for it, forming 33% out of the total sum. Then, numerically questioning, the "Not at all" value was chosen by 12 practitioners, forming a portion of 10% of the population, But the positive value "To a greater extent" had been chosen by only 4 respondents, making up 4% of the participants. To sum it up, « To some extent » and « To the least extent » values were the major judgement values that were mainly opted for by the respondents.

#### **4.17 As An Agent In The Field, Do You Think That The Writing Section Of The Exam Meets Your Expectations?**

This question had generated a large portion of partially satisfied respondents with the National Baccalaureate Exam writing construct, as 79 respondents opted for the "Partly" value, constituting a percentage of 67% out of the total population. Then, significantly, 22 respondents had ticked the "Not at all" value making up 19% of the total participants. The least number was ascribed to the "Totally" value as 17 respondents went for it, which made 14% out of the 118 participants. To sum it up, participants who chose the "Partly" judgement value were the dominant, followed by those who were not at all happy with the current state of the Exam.

#### **4.18 Participants' Comments About The Topic Of The Investigation And Measures That Ought To Be Initiated**

This item raised various concerns of the participants, which would be clustered thematically according to their mutuality and degree of convergence. The table below summarizes and abridges all the resulting comments and concerns of EFL practitioners in Morocco with regard to the topic of this investigation.



Participants' comments	Concerns
<ul style="list-style-type: none"> <li>➤ Writing is not given much importance as it is ascribed 10/40 in comparison with other components: Comprehension and Language.</li> <li>➤ The writing construct is always found to be poorly designed, and the space provided does not suffice for a well-constructed piece of writing.</li> <li>➤ Sometimes, the writing tasks don't respond to the testees' needs. Also, the writing tasks don't target the communicative skills of students.</li> </ul>	<ul style="list-style-type: none"> <li>➤ A meagre value is given to the writing construct in the Exam.</li> <li>➤ Poor design of the Writing Construct.</li> </ul>
<ul style="list-style-type: none"> <li>➤ Mismatch between what is advocated in classroom writing practices and the assessment/evaluation of the writing skills in the Baccalaureate Exam.</li> <li>➤ The writing tasks have to be varied and not repetitive.</li> <li>➤ We need to update the whole Exam parts and rubrics.</li> <li>➤ I think writing and other skills in our Moroccan syllabus are still dealt with as exam oriented skills, which can never help achieve their communicative aims. Testing in Morocco needs reconsideration.</li> </ul>	<ul style="list-style-type: none"> <li>➤ Mismatch between theory and practice.</li> <li>➤ Communicative skills in writing are not given suffice room.</li> </ul>
<ul style="list-style-type: none"> <li>➤ I suggest first to review and update the ELT Guidelines following the world educational revolution.</li> </ul>	<ul style="list-style-type: none"> <li>➤ Revision of the ELT Guidelines in Morocco is a live need.</li> </ul>

As gleaned from the above table, there was a high frequency rate of practitioners who shared worries about the current state of the National Baccalaureate Exam of English as a whole and chiefly the Writing Construct as they advanced the idea of immediate reconsideration of its design as well as its content.

In conclusion, the aforementioned findings revealed an array of facts about the usefulness of the National Baccalaureate Exam Writing Construct, which needs to be paid more than just lip service. Hence, the discussion section would explore the findings in a critical way so as to fortify the strength, validity and reliability of the claims and conclusions that are to be made later. In other words, the evaluation grid and the google form questionnaire, as well as the researcher's readings and unstructured talks with in-service EFL teachers and ELT supervisors, had geared the investigation's course towards a concise discussion of the matter, albeit the remarkable dearth of investigations of the topic.

## 5. Discussion

The primeval role of any type of testing and evaluation in society is well-known, let alone a high-stakes one (Wall & Alderson, 1993; Cheng, 2001). Hence, it is worth mentioning that regardless of its sensitivity, especially with accountability issues to the stakeholders and educational authorities, the social dimensions of the National Baccalaureate Exam made it a fertile territory for investigation, as it was conveyed by the comments of the 118 participants in the study. By the same token, since tests should be judged not only by their technical efficiency but also by whether they were educationally profitable, as Wiseman (1961) argued, the discussion mode adopts a bi-modal way of analysis and synthesis in order to gather between, analogically, convergent results of both the evaluation grid and the questionnaire.

Without further ado, the findings section resulted in many pertinent facts about the topic under investigation, revealing that the topic bears unequivocal interrogation marks that tangible answers are to be provided so as to quench the thirst of EFL/ELT practitioners in the field.

### 5.1 The Theoretical Potency Of The National Baccalaureate Exam Writing Construct

Findings from this study show that apparently, the National Baccalaureate Exam Writing Construct respects to a reasonable extent the standards of a well-structured test once it comes to practicality and familiarity of tasks, albeit the incongruence of topics with regard to the target stream's educational orientation; a deficiency in terms of the unsuitability of tasks with the scientific stream students' intents was pinpointed through the usage of the evaluation grid. In other words, such a frightening discrepancy exists between the construct's tasks and the students' streams or specialty in terms of the degree of challenge that testees face in most tasks over the investigated years, which does match neither their expectations nor the teachers' insofar as the test tasks' challenge is below Baccalaureate students' level or in Krashen's term I-1 (Input minus one) instead of I+1 (Input plus one). Analogically vital,

the cognitive domain in the studied samples was at stake in contrast with what was enunciated in the Standard-based approach that embellish the 2007 thus far Guidelines of Teaching English in Morocco, as well as Bloom's taxonomy, which sketch out the cognitive trajectory of a language learner from low to high-order thinking levels that divert from remembering and understanding to convert in applying, analyzing, evaluating and creating stages. Therefore, if we take the test usefulness for granted without any least consideration of the examinees' competencies on the one hand and the congruence of the exam's tasks with the theoretical guidelines that had been set forth for the purpose of yielding a useful test on the other hand, we will, then, fall in the trap of producing empty minds with faked scores as it was advanced by various authors in other thematically adjacent contexts Gronlund (1977), and Miller, Linn, & Gronlund (2009). As a matter of analogous illustration, Naoua (2016) had conducted a similar study in Algeria evaluating tests administered to Technology Streams from 2001 to 2006, and he concluded that the problems with the tests were due to both theoretical and procedural deficiencies that stakeholders tend to overlook as long as no seriously sharpened complains are raged against them.

Ornately scripted, the Pedagogical Guidelines of Teaching English in Moroccan high schools seem to lay solid grounds for it. As an illustration, it is stated that *"In the dynamic process of writing, a fluent writer resorts to a great deal of sub-skills and strategies simultaneously, in order to make her/his writing effective at the level of rhetorical organization, linguistic variables, purpose, and audience"* (The Pedagogical Guidelines of Teaching English, 2007, p. 27). Unfortunately, the grounds that were theoretically set forth turn to be vulnerable once the usefulness of educational and pedagogical practices is put into question, among which the Baccalaureate Writing construct usefulness is just a case in point. Accordingly, the findings of both instruments of this investigation reveal that educational accountability is at stake whenever the concerned party is the assessment authorities owing to a bulky percentage of teachers and supervisors who took part in this study went for the claim that concerted efforts are to be done in order to hone the usefulness of all of the National Baccalaureate Exam constructs not solely the writing one. Moreover, and conspicuously appealing, some participants even complained through their comments by the end of the questionnaire about the grading scale of the Exam.

To exemplify it, a frequently made interrogation is "Why is the writing construct ascribed only a score of 10/40 within the whole Exam? Is not it important? Is not it an encompassing skill?", three questions that overtly state the fact of holding accountable those who are in charge of designing such a high-stakes exam.

Similarly important, previous studies in the Arab world to state but a few; Kentab (2016), Bedmar (2014), Karami (2013), Almajal (2009), Alkharaif (2008), and Alharthi (2008) as well as studies in the MENA countries, that variously approached EFL writing assessment and evaluation in different areas of the region, such as Ahmed (2010), Gebril and Brown (2014), and Hargreaves (2001) in Egypt. Then, Hamzaoui-Elachachi (2010) in Algeria, Athimni (2017), Barakaoui (2007), Ghrib (2001), Daoud (1999), and Mahfoudhi (1998) in Tunisia. Hence, regardless of the territorial distance similarity of problems, when it comes to inconsistencies of EFL writing assessment and evaluation in the public sector, unify all of the aforementioned countries, unfortunately. By the same token, the 2007 World Bank Report, which is one of the most institutionalized investigations, had unveiled an amalgam of misconducts of authorities, mismatches between official documents and their real realizations, governance issues, and then the irresponsibility of stakeholders in many areas in their countries without marginalizing the educational one.

The WBR report is, to a large extent, in congruence with the findings of the meanwhile investigation concerning the discrepancy between the theoretical foundations and their implementations. Interestingly, the findings of the first part of the evaluation grid, particularly items 8 through 12 of the questionnaire, corroborate with the results sieved from symmetrical investigations, mainly Melouk (2001), who, around twenty years ago, highlighted that evaluation is its initial stages in Morocco because of five factors which would be summarized as it follows:

First, teachers and supervisors lack training in the technical aspects of evaluation. Second, evaluation field research is scarce in Morocco. Third, an independent body responsible for examination, evaluation, and production does not, practically exist in Morocco, not mentioning the National Center of Evaluation, Examinations, and orientation. Fourth, the teaching and learning milieu in which national examinations are run makes teachers' incorporation of continuous assessment in their daily practices a laborious matter, and finally, teachers share a common belief that they are not concerned with evaluation and testing issues, an assumption that is proved by the findings from the current study.

As a consequence, this compatibility between the findings of a study that was carried out in 2001 and another one that was conducted in 2019 is not a mere coincidence but a serious incongruence that must be heeded more than just talks and conference proceeding papers if the educational system is to be saved otherwise the chasm between the higher educational authorities and the hands-on practitioners namely teachers and active field supervisors is being broaden.

Consistently, the outcome of the twelfth item of the questionnaire revealed that a bulky, if not a major majority of respondents ascertained the partial correspondence between the actual realization of the writing construct of the exam and the theoretical Guidelines of Teaching English in Moroccan high schools, Bouziane (2017) studied the practices of EFL reading and writing assessment in Morocco and came out with strikingly annoying findings their state. Notably, among the striking findings that were yielded by Bouziane's investigation is that assessment research of EFL reading and writing in Morocco is scarce and that writing assessment has been generally unfair and harmful to students during and after schools.

It should be made clear that the studies that are cited here were the only ones that the researcher could find in the domain of EFL testing and assessment regarding the topic of evaluating EFL writing assessment practices in Morocco. Nonetheless, the international EFL literature bank is replete with studies that target similar issues but Morocco. Significantly surprising, a strongly warranted convergence between the outcome of the evaluation grid and the reactions of respondents from the online questionnaire vis-a-vis the pragmatic presence of the five standards or the 5Cs of the Standard-Based approach, which is a teaching vision based on Communication, Culture, Comparison, Connection and Community standards that must be inextricably treated, in the investigated construct upon which the educational system is boastfully claimed to be built on.

The results showed that only some standards were present in all samples, namely Connection, Comparison, and indirectly Communication. However, a weird absence of the standard of Culture was greatly obvious in Scientific and Literature stream samples alike; a remark that questions the compliance of the National Baccalaureate Exam Writing construct with the orientations of the guiding Guidelines of EFL Teaching in Moroccan high schools, ministerial circulars and the curricular objectives on the one hand and the aims of the Moroccan constitution which focuses on the integration of not only the local but the global culture as well on the other (Moroccan Constitution, 2011, pp. 16 -17).

Equally important, the findings of this study run again parallel with the ones that were resulted from Melouk's investigation in 2001 about the status of EFL in Morocco, wherein he sifted through teachers' opinions about the National Baccalaureate Exam, then. To his surprise, and presumably ours too, he concluded that teachers in their great majority were unhappy about the state of EFL evaluation in Morocco as their feeling was related not merely to the caveats of current practices then but especially to a concern with the pedagogical value of testing and evaluation.

In a nutshell, the mismatch or partial correspondence between the actual National Baccalaureate Exam is supported by other precedent studies about EFL testing and evaluation in Morocco, the MENA and Arab world countries mainly Bouziane (2017), and Amrous and Nejmaoui (2017) in Morocco, Al-Issaei (2012) and Al-Hajri (2014) in Oman, Sharawneh (2012) and Sobeh (2007) in Palestine, Shukri (2014), Mohammad and Hazarika (2016) in Saudi Arabia, and finally, Alfaki (2015), Zakaria and Mugaddam (2013) in Sudan. Concisely put, as an answer to whether the writing construct complies with the theoretical EFL Guidelines in high schools and the organizational circulars, mainly 2010 and 2014 or not, partial compliance was detected via the evaluation grid and the respondents' reactions to the question that was asked to unravel the intricacies of this point. Therefore, the test blue prints that are to be respected and abided by seemed to be of relative if not trivial importance in the National Baccalaureate Exam writing construct tasks in the studied samples in comparison with what is theoretically ratified and recognized in international models of language testing and evaluation such as Bachman and Alderson's (2000, 2006), Luoma's (2004), Purpura (2004) and O'Sullivan (2012).

### **5.2 The National Baccalaureate Exam Writing Construct And Efl Practitioners' Satisfaction**

The obtained results displayed negative answers to the infused question "Are practitioners and hands-on professionals satisfied with the writing construct of the Moroccan National Baccalaureate Exam?" from both the evaluation grid and the participants' reactions to the online questionnaire as a great deal of concern was raged against the unsatisfactory content and design of the writing construct.

The evaluation grid showed that replication of the same writing types, mainly paragraph, and informal e-mail, as well as the unjustified repetition of well-exhausted genres of writing in consecutive years for both streams, as was shown in the result section for precision matters. So, due to the scarcity, if not the rarity, of evaluative research on the Baccalaureate Exam in Morocco, it turns out to be helpful, if not obligatory, to consult similar studies such as Akpınar and Cakildere (2013) in Turkey, Al Seyabi and Tuzlukova (2014) in Sultanate of Oman, Muhamad and Kiely (2018) in Malaysia and then Ahmad and Rao (2012) in China; these studies albeit the contextual divergence investigated the impact of various End of Term Exams that most of instances share similar pitfalls and shortcomings.

Notably remarkable, all of these studies showed divers degrees of dissatisfaction among EFL practitioners with the investigated Exams. In other terms, the plight of this overarching dissatisfaction shown through the results of the questionnaire above is not a new one, but it is almost a recurrent situation since worries about the state of EFL evaluation and testing were warned against

years ago as they keep prevailing each year in the correction centres of the National Baccalaureate Exam between EFL teachers and ELT supervisors, delegation meetings, academy conventions, and ministry press debriefing conferences but they are sooner cast away in the shadow of underlying backstage political battles between various opponents who seem to enjoy monkey work under the Press spotlight in eventual Press conferences.

Bluntly stated, the *status quo* of many EFL practices, not just testing and evaluation, is madly stagnant on hierarchical scales not just in Morocco but in a wide range of Third World countries (Bedmar, 2014, p. 97; Nkosana, 2013, p. 1; World Bank Report, 2007, pp. 17-20). With reference to the findings of this study, the dissatisfaction is instigated not only by what has been mentioned before but also by the kinds of prompts that were given to stimulate students' brains in stark contradistinction with what is advanced in the National Chart of Education and the newly Strategic Vision of Reform 2015-2030 about integrating multi-techniques of brainstorming students' ideas. Unfortunately, there is extensive unexplained usage of statements all the time in the investigated samples, forgetting about learning styles and individual differences between students.

To illustrate the aforementioned, the survey resulted in a freaking ratio about the inappropriateness of tasks to all examinees as well as the absence of individual differences' consideration in the proscribed tasks in the sense that analogical templates or modals of the exam tasks were frequently used every year making students dull and liable to degradation in their skills and competencies. Similarly important, because of what has been mentioned, testees, lately, appear to come to the Exam with only two or four templates in mind, a hurdle that is going to humper their career later on. Thereof, the effects of such exams are detrimental, as was proven by various washback or backwash -in Hughes' terms- studies in other parts of the world, such as Nazari and Nikooupour (2011) in Iran, wherein the emphasis of the study was focalised on the effect of High School Examinations on Iranian EFL learners' language learning beliefs, Sultana (2015) in Bangladesh who studied the usefulness of Higher Secondary Certificate (H.S.C) examination and its format impact on testees' scores and Naoua (2016) in Algeria who evaluated the reliability and validity of administered Tests to Technology Streams. Statistically significant, the respondents' ratio is much alluring when authenticity matters are concerned, as the findings indicated that more than half of the practitioners were unhappy with the presence of tasks that don't match the real world, an issue that must, normally, be addressed with more zeal and caution but the analysed samples conveyed the opposite of what ought to be. This bitter reality results from the ignorance or negligence, be it overt or covert, of the main agents in education or what was called in the 2007 Wold Bank Report "Main agents in the education process" (p.11). Therefore, incentives, be it emotional or pecuniary, are to be granted to teachers and ELT field supervisors as well as take their opinions and suggestions about the major concerns of education, such as the National Baccalaureate Exam writing construct usefulness into account with a real sense of responsibility and a high degree of accountability (Kentab, 2016; Wold Bank Report, 2007). In addition to all of what has been said and stated earlier, a deeply frank and purposeful dialogue must take place between the specialists in testing, the designers of the exam, the primate EFL/ELT practitioners, teachers and field supervisors, educational activists and the educational authorities to discuss the problem of mismatching between theory and practice of the concerned issue and therefore bridge the gap between all of the concerned agents on the one hand and safeguard the theoretical potency of the National Baccalaureate Exam on the other (Tawil, Cerbelle, & Alama, 2010, pp. 85-86).

### **5.3 Fairness and Equity of The National Baccalaureate Exam Writing Construct**

Eye catching are the findings that were resulted from the survey items that addressed the issues of fairness in terms of grading scales, individual differences, and high-order thinking skills. Similarly interesting, the comments of participants by the end of the survey resulted in a notable disclosure of worries about the current state of not just the writing construct of the exam but the entire entity of the exam as a whole.

With regard to the findings about the scaling rubrics and scoring schemes that were used in the studied construct, almost half of the participants agreed that they were well-explained. However, another group of participants who represented more than two thirds went to the other extreme, stating that the scoring scales and schemes are not well-explained. Therefore, the perceived contrast between the two groups shows that there is a hidden misunderstanding of scoring and grading scales from the participants' side because if we snap a look back at the evaluation grid's results, it conveys that the scoring scales were not explained well. As an illustration, both of the evaluation grids that were adopted presumably showed that the grading scales were not clear enough in the case of samples that used only one task, namely the 2014 National Baccalaureate Exam Writing construct (Literature and Scientific stream), a holistic score scale of 10/40 was adopted without any further analytical clarification for the examinees then from 2015 hitherto a bi-scoring scale of 10/40 has been adopted with 4/10 most of the time for guided tasks whereas 6/10 for the semi-guided or free tasks. Consequently, this result questions not just the test designers but the respondents' assessment literacy as well because of the higher rate of agree or disagree values of the two camps; pros and cons of the grading scales and schemes. Hence, practitioners' assessment literacy should be put to the test too by going back to sift through their pre-service training modules wherein a meagre portion is ascribed to testing and evaluation that is, in most instances, done in a helter-skelter manner at the end of the training period.

O'Sullivan (2012) had argued that engineering and rating problems in language testing are due to a shortage in teachers' and test designers' assessment literacy on the ground that most of them imitate others' practices if not rely on individualistic judgement and common sense evaluation. So, the problem of scoring scales absurdity would be a case in point that needs to be investigated aside by the educational authorities.

Heading back to the issue of the degree of higher-order thinking skills and presence of individual differences, the survey and the evaluation grid displayed negative results about their presence status in the concerned construct tasks. Higher-order thinking skills were given much roaring utility since the era of Benjamin Bloom with his taxonomy of objectives that learners must attain at different but gradual phases or stages (Santrock, 2011, p.403). Bloom's taxonomy has been used by many teachers, educators, and educational systems in planning and creating learning goals and objectives due to its well-grounded standards (Bart, 2008; Lord & Baviskar, 2007). In its first version, Bloom's Taxonomy encompassed six concepts are Knowledge, Comprehension, Application, Analysis, Synthesis, and Evaluation before it was innovated by Alderson, who changed the labels from noun phrases to verb phrases due to the fact that thinking skills are not still but active.

As mentioned earlier, with the advent of the Standard-Based approach, the Moroccan educational system adopted Bloom's Taxonomy in various ways and shapes, mainly in lesson planning and pedagogical activities as the 2007 Guidelines of EFL Teaching in High Schools (pp.30-50). However, the validity of the realization of the theoretical tenets of EFL frameworks, including higher-order thinking skills' principles, in Morocco seemed to be wane if not at stake, according to the meanwhile study findings.

Accordingly, two focal matters thread their ways through the discussion of these findings: the concept of accountability and governance in relation to the participants' expectations that were not fully met in the samples under investigation because of the obvious replicable nature of the construct tasks along with other issues that were early approached. Likewise, the majority of participants are outraged by the fact that they strive the whole year to teach certain genres or models of writing without even heeding attention to the time that is consumed in doing so, but to their disappointment, the exam tasks turn to be banal and of a low-thinking skills nature.

Remarkably, some participants, as displayed in the comments table above, asked for a complete revision and rejuvenation of the outdated standards and principles of EFL teaching in Morocco, starting with the 2007 Guidelines of Teaching English in High Schools, syllabi, modes of teaching and, decisively, testing measures among which we find the Baccalaureate Exit Exam as a high-stakes one, that has detrimental effects on the examinees' lives and social prosperity, therefore.

As matters of deficiencies in terms of scoring scales and equity in the National Baccalaureate Exam writing construct have emerged and, therefore, questioned in this investigation, so too have the issues of accountability and fair governance in the educational system which was alluded to beforehand. Hence, Cizek (2001, 2003) and Mehrens (1998) (as cited in Ghaicha, 2016, p. 217) had argued that students', in this case, Baccalaureate students, achievement in standardized and norm-referenced tests would insure accountability of teachers and schools. However, this reductionist view is countered here by the fact that the design and content of the Exam itself are questionable with respect to the results of this investigation, as the teaching and pedagogical staff could not be held accountable before the administration and execution of the designed exam, which is in fact a field of lethal combinations of worries that have to question its engineering agents in the corridors of the ministry. Therefore, fair governance and just accountability measures demand a frank bilateral as well as multi-lateral and horizontally based dialogue to yield good results that will safeguard not just the usefulness of a high-stakes exam such as the National Baccalaureate one but the productivity of the educational system in its integrity as well. Equally vital, an evaluation study was conducted by Kentab (2016) in Saudi Arabia about the presence of accountability and fair governance in the Saudi educational system, stating four main barriers to accountability in the system (p.149):

- a. Lack of technically qualified human cadres and personnel for the implementation of accountability measures.
- b. The rarity of training programs that are directed to the implementation of accountability in work
- c. Limited availability of appropriate measures to measure performance.
- d. The low level of awareness of the role of the media in accountability is important.

By the same token, the Moroccan educational system would be liable to the above stated obstacles, taking into account a variety of symmetrical patterns between the Moroccan system or regime and the Saudi one. Consequently, the shortcomings of the National Baccalaureate Exam writing construct are due to certain practices that need to be addressed seriously in light of specialized and empirically independent knowledge rather than in the blurred shadows of political agendas.

On the aggregate, it is rudimentary urgent that the shortcomings of the National Baccalaureate Exam writing construct be investigated and discussed in a socially, pedagogically, and educationally open debate in order to make up for any deficiency or

mismatch between theory and practice, failure of meeting the educational interventionists' expectations, the weird shyness of the presence of individual differences and higher-order thinking skills in the writing tasks attested through the investigated samples, the absence of culturally of moderated topics and themes in the ascribed tasks, the lame presence of the standards of the Standard-Based approach, the feeble explanation of the scoring scales and schemes to examinees, and last but not least the flagrant and unjustified recurrence of particular types and genres through consecutive years. So, without further wording, the above stated shortcomings are the most decisive ones that this investigation had revealed as some of them are in congruence with Melouk's study results (2001) about EFL testing and assessment in Morocco as well as the results of the World Bank Report in 2007 and Bedmar's too about educational reforms in Morocco (2014).

In brief, the Moroccan National Baccalaureate Exam writing construct must be heeded with sincere attention so as to fix its pitfalls and strengthen its assets, and therefore, give it due importance in the testing and evaluation spectrum in Morocco.

## **6. Conclusions and Recommendations**

The study substantially succeeded in evaluating the concerned construct in an objective-like manner through both analyzing hard paper samples of the Baccalaureate Exam writing construct over the labelled time period earlier and the EFL teachers' as well as ELT supervisors' opinions about the issue via the usage of an online displayed survey.

Accordingly, the three objectives of the study were fully met despite the difficulty of analyzing every point or matter that could shape the course and results of the investigation. However, only one of the objectives, which was the second one, was only partially attained for some reasons that are to be stated in due time. In addition, reliability and validity measures of results were neatly scrutinized to shun any potential misinterpretation or misjudgment.

The researcher, therefore, kept one's guard against being too subjective or intrusive despite his position as a practicing EFL teacher in a public high school.

Significantly attracting, the research results support what was claimed by Melouk (2001) concerning the blurred nature of assessment and testing practices in Morocco back then. Thus, though the vast span of time between the two studies, the dwelling pitfalls and shortcomings of assessment and testing practices in Morocco bridge the time gap between the two studies. Hence, in light of this impressive but expressive compatibility between the two studies, regardless of their time era divergence, the objectives of this study will be stated according to their fulfilment order in the main course of the paper.

To demystify the conclusions of this study, its first objective was fulfilled in a two-folded way, which was adopted to infiltrate evidence that would either help or hinder the course of this research. Therefore, a body of literature about testing, constructing, and scoring a given language test was consulted, mainly the theoretical frameworks set forth by Bachman and Palmer (1990, 1996, 2002), Luoma (2004), and O'Sullivan (2012), to state but a few.

Then, this objective was empirically fulfilled through the usage of evaluation grids to analyze exam samples of both Literature and Scientific streams with a critical eye on the basics of a sound design of a language test that were already discussed. Similarly, the online questionnaire was another way that fortified the attainment of the first objective by digging deeper into the minds of the chief interventionists in the field, who were all seriously eager to participate and opine about the issue.

As the foregoing exposition should make clear, the second objective investigation was partially met using the pre-mentioned means of investigation to elicit the practitioners' stances about the issues that were provoked in the questionnaire. To our surprise, the answer to reach it turned out to be negative in the sense that the design, the task format and content, and then the type and genre of writing were not given enough attention, which made them void of theoretical solidity. However, the partial fulfillment of this objective stems from the fact that the researcher had access to only a very modest portion of studies and papers that critically studied this construct by comparing its principles with other widely recognized testing frameworks.

With regard to the last two objectives of this study, they were greatly met due to the researcher's reading, analysis, and synthesis, along with the constructive and practical comments of the participants that made the task of postulating novel ideas about how to better the quality of the National Baccalaureate Exam writing construct a feasible matter in the sense that many ideas were inspired by the researcher from different stations during the journey of this study. Moreover, the theoretical deficiencies that were detected during this investigation will imperatively highlight the way EFL practitioners in Morocco and enhance the productivity of the Moroccan educational system. Nevertheless, it would be wise stating that this study tried its best with all the available means to bring to light such a sensitive topic which concerns not merely the researcher himself but the future of a whole nation as well.

So, collective contribution and earnest engagement of all citizens are highly requested in this globalized world so that the nation's ship would safely dock on the golden shores of this era regardless of the current long-lasting educational turmoil between the Moroccan Ministry of Education and teachers over the 2023-2024 New Basic Law of Practitioners in Education in Morocco that led to fiercely long strikes for more than three consecutive months.

### 7. Pedagogical Implications

The current investigation has unraveled some of the inconsistencies, as well as deficiencies of the National Baccalaureate Exam writing, construct in a substantially objective way by adopting scientific research techniques and basics. Consequently, a multitude of positive and negative points have been detected throughout this journey, which resulted in the following recommendations or tips to the concerned EFL stakeholders, National Agency of Evaluation and Examinations, ELT supervisors, and EFL teachers in order to work together towards enhancing the usefulness of the concerned construct as well as hone our students' skills:

1. An urgent revision of the theoretical foundations of the writing construct must be done on the basis of well-recognized models such as Bachman and Palmer (1996), Banergee and Luoma (1997), Gallagher (1998), and Weigle (2002).
2. Higher-order thinking skills must be granted more than just lip service by incorporating tasks that provoke examinees' thinking skills and avoid elementary writing tasks that favor nothing more than low-order thinking skills, if not mental retardation.
3. The EFL Guidelines (2007) in high schools are ought to be rejuvenated to suit the requirements of the new era as well as fulfill the ambitious aims of the Moroccan Strategic Vision of Education 2015-2030.
4. The standards of the standard-based approach should be respected by including them in the writing tasks.
5. Culture as a parcel of values must be considered in the writing tasks.
6. Individual differences are to be considered through designing tasks that target both female and male interests.
7. A variety of writing types and genres exist in English. Therefore, there must be a variety of writing types and genres in the writing construct rather than repeating the same types and genres, mainly informal e-mails and short paragraphs, which may create dull and lazy students.
8. The writing tasks difficulty should logically be challenging to Baccalaureate students. Students and teachers feel victimized hustling the whole year with highly thought-provoking and timely consuming topics to find common core or false beginners' topics in the Baccalaureate Exam writing construct.
9. The writing skill is an intricate skill as it encompasses a variety of competencies. Thus, it must be given at least an equal scoring scheme/rubric as Language and Comprehension constructs instead of only 10/40.
10. One of the writing construct tasks is ought to be totally free, not even semi-guided, so that examinees would write freely and inspirationally.
11. The scoring rubrics must be clarified and stated clearly hand-by-hand with the tasks; there should be an analytical scale instead of a holistic one, at least for testees.
12. EFL assessment literacy must be given due attention; otherwise, evaluation and testing chaos is on the horizons.
13. Writing tasks' themes/topics must be congruent with the examinees' specialty or stream.
14. With a critical eye, imitate other successful testing frameworks of other countries like the European Common Framework of testing and evaluation, but the guard must be kept against any kind of blind imitation.

Finally, it would be extremely imprudent for a democratically driven nation to keep ignoring research findings and recommendations in all domains, not just in education. Therefore, the above recommendations are not binding, compelling, or even offending, in case it might be conceived in such a way, but they were stated to serve the scientific requirements of the study and the benefits of Morocco's future citizens alike.

By the same token, it is noteworthy confessing that the pre-mentioned recommendations are not the only ones that this study yielded, but we reined the self to state the most practical, pertinent, precise, and possible ones in Morocco. As a consequence, further research is needed to probe deeper into the unperceived aspects, ramifications, and repercussions of the assessment and testing of EFL in Morocco. In other terms, concerted and relentless efforts must be done by academic researchers to unveil the misdemeanors as well as the maladministration deeds of the educational and political backbenchers' policies. We are, therefore, confident that our research will serve as theoretical scaffolding for future studies in the realm of EFL assessment and testing, among which the current investigated topic is just the tip of the iceberg in a stormy ocean of multifarious educational, political, and even ideological agendas and plans.

In a nutshell, as far as serious and honest researchers exist, probing questions will ever exist. So, we are greatly hopeful that more studies will be conducted in this area though the bleak educational future that is lurking around our nation's corners.

**Funding:** This research received no external funding.

**Conflicts of Interest:** The authors declare no conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

## References

- [1] Abulalaa, S. (2013). Educational accountability as realized by teaching staff members in Education College at Aswan. *College of Education, Culture and Development Magazine*, 72, 45-117.
- [2] Agence Marocaine de Coopération Internationale. (2011). *Formation au Maroc/ Formation. AMCI professionnelle*. Retrieved from <http://www.amci.ma/formationProfessionnelle.asp?r=51&sr=57>
- [3] Ahmad, K., Ghani, M., Alam, M., & Sadiq, M. (2012). Washback effects on English language teachers. *International Researcher*, 1 (3), 181-190.
- [4] Ahmed, H. (2006). Accountability and its relation with school performance quality evaluation from planning perspective, Cairo, *Arabic Education Future Magazine*, 12, 9 -150.
- [5] Akapinar, K. and Kalkidere, B. (2013). Washback effects of high-stakes language tests of Turkey (KPDS and UDS) on productive and receptive skills of academic personnel. *Journal of Language and Linguistic Studies*, 9(2), 81-94.
- [6] Alakayleh, A. (2017). Could different items arrangements affect 10th grade students' performance in multiple-choice tests in maths, science, and English language final exams? *Journal of Education and practice*, [Online] 8(17), 186. Available at: <http://www.iiste.org> [Accessed 20 Jun. 2019].
- [7] Alamri, H., & Barakat, M. (2004). *The reality of educational accountability in Ministry of Education in Jordan: an analytical developmental study*. (Unpublished PhD. Dissertation), Jordanian University: Amman, Jordan.
- [8] Alderson, C., & Hamp-Lyons, L. (1996). TOEFL preparation courses: A study of washback. *Language Testing*, 13 (3), 280-297. doi:10.1177/026553229601300304.
- [9] Alderson, C., Clapham, C. & Wall, D. (1995). *Language test construction and evaluation*. Cambridge: Cambridge University Press.
- [10] Alderson, J., & Wall, D. (1993). "Does washback exit?" *Applied Linguistics*, 14, 115-129.
- [11] Al-Jamal, D., & Ghadi, N. (2008). "English language general secondary certificate examination washback in Jordan." *The Asian EFL Journal*, 10 (3), 158-186.
- [12] Amos, I. (2014). Formative and summative assessment in educational enterprise. *Journal Of Education And Practice*, 5(20), 111-112-113.
- [13] Amrous, N., & Nejmaoui, N. (2017). *A developmental approach to the use of critical thinking skills in writing: The case of Moroccan EFL university students*. ASELS Annual Conference Proceedings, Mohammed V University of Rabat, Morocco.
- [14] Ana, P., & Marta, E. (2010). Washback of an oral assessment system in the EFL classroom. *Language Testing*, 27, 33-49.
- [15] Andrews, S. (1994). The Wash back effect of the examination: its impact upon curriculum innovation in English language teaching. *Curriculum Forum*, 44-58.
- [16] Babae, H. (2017). A snap shot on construct validity in language testing: definitions and implications. *Journal of Education And Practice*, 8(18), 119.
- [17] Bachman, L., & Palmer, A. (1996). *Language testing in practice: developing and designing. Useful Language Tests*. Oxford, England: Oxford University Press.
- [18] Bachman, L. F. (1990). *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.
- [19] Bailey, K. M. (1996). Working for washback: A review of the washback concept in language testing. *Language Testing*, 13(3), 257-279.
- [20] Bailey, K. M. (1999). *Washback in language testing*. Princeton, NJ: Educational Testing Service.
- [21] Banque Mondiale. (1995). *Rapport de la Banque Mondiale de 1994: dépense publique au Maroc*. Rabat: Banque Mondiale.
- [22] Bany Yaseen, O. (2019). Educational assessment modern strategies. *Journal of Education And Practice*, [online] 4(1). Available at: <http://www.iiste.org> [Accessed 20 Jun. 2019].
- [23] Barakaoui, K. (2007). Rating scale impact on EFL essay marking: A mixed method study. *Assessing Writing*, 12, 86 –107.
- [24] Bedmar, L.V. (2014). Educational reforms in Morocco: Evolution and current status. *International Education Studies*, 7(12), 95-105.
- [25] Belfkih, M. (2000). *Appearing in the program Hadith Al-Arbiâ. Broadcast on the 2N television channel Wednesday, October 4*.
- [26] Berry, V., & Lewkowicz, J. (2000). Exit-tests: Is there an alternative? *Hong Kong Journal of Applied Linguistics*, 5(1), 19-49.
- [27] Boutalet, H. (1996). *Réflexions sur l'enseignement supérieur et la recherche scientifique au Maroc*. Casablanca: Éditions Maghrébines.
- [28] Bouziane, A. (2017). Why should the assessment of literacy in Morocco be revisited? In S. Hidri & C. Coombe (Eds.), *Evaluation in foreign language education in the Middle East and North Africa* (pp. 305–314). New York: Springer International Publishing.
- [29] Buck, G. (1988). Testing listening comprehension in Japanese university entrance examinations, *JALT*, 10 (1), 15.
- [30] Cheng, L. (2005). *Changing language teaching through language testing*. United Kingdom: Cambridge University Press.
- [31] Cheng, L. (1997). How does washback influence teaching? Implications for Hong Kong. *Language and Education*, 11(1), 38-54.
- [32] Cheng, L. (1999) Changing assessment: Washback on teacher perceptions and actions. In *Teaching and Teacher Education*, 15(3), 253-271.
- [33] Cheng, L. (2003). Looking at the impact of a public examination change on secondary classroom teaching: A Hong Kong case study. *Journal of Classroom Interaction*, 38(1), 1-10.
- [34] Cheng, L. (1998). Impact of a public English examination change on students' perceptions and attitudes towards their English learning. *Studies in Educational Evaluation*, 24 (3), 279–301.
- [35] Commission Spéciale Éducation Formation, Royaume du Maroc. (1999). *Charte Nationale d'Éducation et de Formation*. Royaume du Maroc. Retrieved from [http://www.uh2c.ac.ma/uh2c/loi/charte\\_fr.pdf](http://www.uh2c.ac.ma/uh2c/loi/charte_fr.pdf)
- [36] Coombe, C., Davidson, P., O'Sullivan, B. and Stoyhoff, S. (2012). *The Cambridge Guide to Second Language Assessment*. 1st ed. New York: Cambridge University Press.
- [37] Cooper, P. (1984). *The assesment of writing ability: A review of research*. No 82-15. Princeton: Educational Testing Service.



- [38] Creswell, J. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th Ed.). Boston, MA: Pearson.
- [39] Creswell, J. (2014). *Research design: Qualitative and quantitative approaches*. Thousand Oaks, California: Sage Publications.
- [40] Davies, A. (2008). Ethics and professionalism. In E. Shohamy & N. H. Hornberger (Eds.), *Encyclopedia of language and education: Vol. 7. Language testing and assessment* (2nd ed., pp. 429 – 443). New York, NY: Springer.
- [41] Décret n° 2-99-924. (1999). *Décret du Chaabane 1420 (3 décembre) fixant les attributions du ministère de l'enseignement secondaire et technique*. Royaume du Maroc.
- [42] Dornyei, Z. (2003). *Questionnaires in second language research construction, administration, and processing*. London: Lawrence Erlbaum Associates, Publishers.
- [43] El-Andaloussi, K. (2005). *50 ans de développement humain au Maroc et perspectives pour 2025*. Retrieved from <http://www.rdh50.ma/fr/pdf/contributions/GT4-1.indd.pdf>
- [44] Gates, S. (1995). Exploring washback from standardized tests. *Language Testing in Japan*, 101-106.
- [45] Ghaicha, A. (2016). Theoretical framework for educational assessment: A synoptic review. *Journal of Education and Practice*, 7(24), 212 – 213.
- [46] Gebriel, A., & Brown, G. T. (2014). The effects of high-stakes examination systems on teacher beliefs: Egyptian teachers' conceptions of assessment. *Assessment in Education: Principles, Policy and Practice*, 21(1), 16–33
- [47] Hamzaoui-Elachachi, H. (2010). Development of a writing curriculum for academic purposes at tertiary level: The case of Algerian EFL university students. *ESP World*, 1(27), 1-9.
- [48] Hanson et al. (2005). Mixed Methods Research Designs in Counseling Psychology. *Journal of Counseling Psychology*, 52(2), p. 224–235.
- [49] Heywood, J. (2000) *Assessment in higher education*. London: Jessica Kingsley.
- [50] Hughes, A. (1989). *Testing for language teachers*. Cambridge: Cambridge University Press.
- [51] Hughes, A. (2003). *Testing for language teachers*. Cambridge: Cambridge University Press.
- [52] Ibrahim, A. (2018). Improving assessment and evaluation skills of public school teachers in Jigawa state, Nigeria. *Journal of Education and Practice*, 9(11), 22.
- [53] Jabbarifar, T. (2009). The importance of classroom assessment and evaluation in educational system. In: *Proceedings of the 2nd International Conference of Teaching and Learning (ICRL, 2009)*, Iran.
- [54] Jacobs, H.L., Zingraf, S. A., Wormuth, D. R., Hartfiel, V. F. and Hughey, J.B. (1981). *Testing ESL composition: a practical approach*. Rowley, Mass: Newbury House.
- [55] Jilani, R. (2009). Problematizing high school certificate exam in pakistan: a washback perspective. *The Reading Matrix*, 9 (2), 175-183.
- [56] Karami, H. (2013). The quest for fairness in language testing, educational research, and evaluation. *An international Journal on Theory and Practice*, 19(2), 158-159.
- [57] Kim, Y., & Lee, J. (2012). Rethinking the standardized test: The pursuit of diversity and standardization. *Procedia- Social and Behavioral Sciences*, 47, 1495-1500. Retrieved from <http://www.sciencedirect.com>
- [58] Knight, P.J. (2002). Summative assessment in higher education: Practices in disarray. *Studies in Higher Education*, 27(3), 2002: 275-286.
- [59] Lazrak, M. & Yechouti, Y. (2017). Issues in Moroccan higher education. *International Journal of English Language & Translation Studies*, 5(2), 86-93.
- [60] Llorente-Bedmar, V. (2014). Educational reforms in Morocco: Evolution and current status. *International Education Studies*, 7(12), 95-105. doi:10.5539/ies.v7n12p95
- [61] Lloyd, D., Davidson, P. and Coombe, C. (2009). *The fundamentals of language assessment*. Dubai: TESOL Arabia.
- [62] Maddolyn, R. (2016). *The impact of high-stakes testing on the learning environment*. (Master's thesis). Retrieved from [http://sophia.stkate.edu/msw\\_papers/658](http://sophia.stkate.edu/msw_papers/658)
- [63] McNamara, T. (2005). 21st century Shibboleth: Language tests, identity, and intergroup conflict. *Language Policy*, 4, 351–370.
- [64] Melouk, M. (2001). The state of EFL evaluation in Morocco: The testers and teachers' opinions. The Teaching and Assessment of English for Global Purposes. Paper presented at *MATE proceedings*, Essaouira, pp. 41–51.
- [65] Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13– 103). Washington DC: The American Council on Education and the National Council on Measurement in Education.
- [66] Messick, S., 1996. Validity and washback in language testing. *Language Testing*, 13 (3), 241- 256.
- [67] Ministry of National Education. (2007). *English language guidelines for secondary schools: Common core, first year, and second year baccalaureate*. Rabat, Morocco.
- [68] Moussaoui, S. (2012). An investigation of the effects of peer evaluation in enhancing Algerian students' writing autonomy and positive affect. *Procedia—Social and Behavioral Sciences*, 69, 1775 - 1784.
- [69] Muhamad, M., & Kiely, R.C. (2018). Understanding Washback in the Classrooms: The Conflicts of the Syllabus and the Examination. *Journal of Education and Practice*, 9 (2).
- [70] Muhamad, M. and Kiely, R. (2018). An Analysis of Focus on Form Practice in Communicative English Language Teaching Classrooms. *IJELTAL (Indonesian Journal of English Language Teaching and Applied Linguistics)*, 2(2).
- [71] Naoua, M. (2016). *An evaluation of English language testing in the baccalaureate examination: The case of the tests administered to technology systems from 2001 to 2006* (Ph.D). Mentouri Brothers University, Constantine.
- [72] Nazari, M. and Jahanbakhsh, N. (2011). The washback effects of high school examinations of Iranian EFL learners' language learning beliefs. *Journal of Language and Translation*, 2(1), 29-49.
- [73] Nourddine, A., & Nejmaoui, N. (2016). A development approach to the use of critical thinking skills in writing: The case of Moroccan EFL university students. *Arab World English Journal*, 142-156.
- [74] Nunan, D. (1992). *Research Methods in Language Learning*. 1st ed. New York: Cambridge University Press.
- [75] Ouakrime, M. (1996). What is high about higher education? The Idea of the University. Rabat: Faculty of Letters and Human Sciences.
- [76] Shepard, L. (1997). "The centrality of test use and consequences for test validity." *Educational Measurement: Issues and Practices*, 16 (2), 5-13.
- [77] Schmidt, P. (2010). *Longman Dictionary of Language Teaching and Applied Linguistics*. Pearson Education Limited.

- [78] Shohamy, E. (1997). Testing methods, testing consequences: Are they ethical? Are they fair? *Language Testing*, 14, 340–349.
- [79] Sobhi, T., Cerbelle, S., & Alama, A. (2010). *Education au Maroc: Analyse du secteur*. Rabat: UNESCO, Bureau multipays pour le Maghreb.
- [80] Tawil, S., Cerbelle, S. and Alama, A. (2010). *Education au Maroc*. Rabat: UNESCO, Bureau Multipays pour le Maghreb.
- [81] The special committee for education and training. (1999) National Charter for education and training. Morocco: Ministry of Education.
- [82] The international bank for reconstruction and development (2007). *The road not traveled education reform in the Middle East and North Africa*. Washington D.C: The world Bank.
- [83] The National Charter of Education and Training, article 96, 97, and 98. (October, 1999)
- [84] The Constitution of the Moroccan Kingdom, the series of the Moroccan Judicial documents, (2011). 43-70-71.
- [85] The Throne's Speech (Feast of Youth). (August, 2018).
- [86] The Strategic Version (2015/2030). 43-70-71.
- [87] Vanderstoep, S. and D. Johnson, D. (2009). *Research methods for everyday life*. San Francisco: John wiley & Sons, inc.
- [88] YouTube. (2019). *Morocco's King Mohammed VI urges government to fix social services*. [online] Available at: <https://www.youtube.com/watch?v=cEnwma-CWsk> [Accessed 21 Jun. 2019].