

BASHAR FARRAN – ILDIKÓ HORTOBÁGYI – GYÖNGYI FÁBIÁN

University of Pannonia
bfarran89@gmail.com
ildiko@almos.uni-pannon.hu
fabian.gyongyi@mftk.uni.pannon.hu

Bashar Farran – Ildikó Hortobágyi – Gyöngyi Fábíán: Testing English as a Foreign Language in Palestine:

A Case Study of INJAZ (GCSE) 2018 English Exam

Alkalmazott Nyelvtudomány, XX. évfolyam, 2020/1. szám

doi:<http://dx.doi.org/10.18460/ANY.2020.1.001>

Testing English as a Foreign Language in Palestine: A Case Study of INJAZ (GCSE) 2018 English Exam

A minőségi nyelvoktatás megkívánja az adott idegen nyelven zajló vizsgák módszertanának, tartalmának, valamint eredményeinek folyamatos felülvizsgálatát és továbbfejlesztését. Ugyanakkor az angol mint idegen nyelv vizsgák nyelvi eredményei Palesztinában (beleértve a Nyugati Partvidéket, Kelet-Jeruzsálemet, valamint a Gázai Övezetet is) nem tartanak számot nagy érdeklődésre az akadémiai körökben. Tanulmányunk a 2018-ban felhasznált angol nyelvi érettségi (INJAZ) vizsgaanyag leíró és analitikus, valamint tartalom-alapú elemzésével azt vizsgálja, hogy a palesztinai középiskolai érettségi vizsga mennyiben felel meg a Közös Európai Keretrendszer (CEFR) elvárásainak. A kiválasztott vizsga olyan tanulók angol nyelvtudását hivatott megmérni, akik tizenkét éven keresztül folytattak tanulmányokat az idegen nyelven. A vizsgaeszköz jóságmutatóit általában, az egyes itemek minőségét pedig a Bachman és Palmer (1996) munkájában bemutatott kommunikatív kompetenciamodell alapján vesszük górcső alá. Az eredmények arra utalnak, hogy az említett érettségi vizsga nemcsak a kommunikatív kompetenciákat nem méri, hanem a Palesztin Oktatási Minisztérium angol nyelv oktatására vonatkozó célrendszerét sem közelíti meg. A vizsgaanyag az írásbeli készségek mérésére korlátozódik, miközben figyelmen kívül hagyja a szóbeli készségeket, amely negatív hatással van mind a nyelvtanulási folyamatokra, mind pedig az angol nyelvi teljesítményekre.

Introduction

Among the various objectives of foreign language assessment, one main aim is to maintain a gateway to enter a modern education institute and advance higher education studies. It also acts as an instrument with which knowledge providers, or teachers, can measure the impact of what has been planned and implemented in their teaching and the learning processes; they can identify the strengths and the weaknesses of the students' competencies, and, as a result, they might be able to propose solutions that emphasize the strengths and assist avoiding the weaknesses or remedy them.

The methods available to complete assessment procedures are varied.

“Assessment can draw information from a wide range of elicitation, observation and data collection procedures, including multiple-choice tests, extended responses such as essays and portfolios, questionnaires and observations.” (Bachman, 2004: 4)

Given the large range of techniques available to elicit students’ language competencies, it is vital for students to be aware of how their efforts are to be assessed, scored, and evaluated, which might motivate them to learn more efficiently, and thus achieve better performance.

English as a foreign language has been taught in mainstream education in Palestine for 12 years. For the 12 consecutive years, students have been following a Macmillan curriculum called *English for Palestine*, which integrates different pedagogical perspectives and aims to develop all the language skills.

Students of Tawjihi (a previous version of INJAZ) in Palestine used to sit for two different exams: a Jordanian one in the West Bank area, and an Egyptian one in the Gaza Strip, while at present they are administered a Palestinian examination. From the 1948 Nakba through 1967 – the time of the Palestinian fight for independence – until 1994 the Jordanian curriculum had been applied in the schools in the West Bank area, and, at the same time, the Egyptian curriculum was followed in the Gaza strip schools (Amara, 2003). In 1994 the Palestinian Curriculum for General Education was planned and was later introduced gradually starting from 2000. Nevertheless, the Palestinian English language curriculum was not implemented until late 2004 (Yamchi, 2006). With the emergence of the newly introduced Palestinian curriculum, the necessity of assessment and evaluation became clear, however, it is merely the assessment of the program implementation. The assessment process gathers information about the efficiency of the attainment of the course aims, about what students have learned, the validity of the objectives, the adequacy of placement and achievement tests, the amount of time allotted for each unit, the appropriacy of the teaching methods and the problems encountered during the course.

INJAZ examination, previously named aka Tawjihi exam, is defined as a final public examination of secondary education and, at the same time, as the entrance examination to university (Nicolai, 2007). Currently, INJAZ examination is claimed to be the general secondary achievement and proficiency test in the state of Palestine and is an essential part of the educational system. It assesses students' performance of English as a foreign language and is administered besides the examinations in the rest of the academic subjects as part of the final examination. The INJAZ examination consists of one test session, except for the Humanities stream (or module), which also has a narrative literature session, thus, depending on the stream

of the study area, the examination consists of different components. While the assessment instrument analysed in this study includes reading comprehension, vocabulary, language, grammar and writing papers, the humanities stream additionally includes a literature section as well.

Designing an INJAZ test is not an easy task. Test developers should be aware of the objectives of teaching, and the topics teachers deal with. Moreover, they should be well informed of some further details of the curriculum content and skilled in compiling tests according to the criteria of a balanced test instrument that handles the difficulty of individual differences and matches the assigned objectives of the syllabi content as well as the levels of cognition. Therefore, the Ministry of Education (MoE) establishes a board responsible for preparing, controlling the process, and assessing the General Certificate of Secondary Education (GCSE) exam papers every year. The goal of this paper is to evaluate the General Certificate of Secondary Examination (GCSE) in Palestine through analyzing the test instrument implemented for the science stream (or module) in 2018.

Overview of English for Palestine for Grade 12 “INJAZ-exam”

English as a Foreign Language in Palestine is mainly introduced through formal education based on a formal Palestinian English curriculum run from grade 1 until grade 12. The Palestinian Curriculum for General Education Report (1996: 12), which establishes the curriculum vision, recommends “teaching English in view of the heavy involvement of the Palestinians with the modern world. It aims to provide a curriculum which enables students to “read, write, speak and appreciate English as a world language by learning it throughout the twelve years of education” (The Palestinian Curriculum for General Education Report, 1996:12). In addition, a comprehensive list of TEFL objectives is set by the MoE in Palestine with a special focus on both the oral and written skills of communication. A further element of the document principles includes the requirements of meeting individual and community needs.

Teaching English used to start from the fifth grade until the year 2000 when the MoE initiated cooperation with MacMillan Education and the current version of “*English for Palestine*” based on British English (RP) was produced. However, despite the revision of the curriculum in every 5 years, the program fails to meet the requirements in many ways (Ramahi, 2018; Bernard & Maître, 2006).

It seems that there is no evidence of integrating an individual psycholinguistic dimension into the curriculum. Another problem of the curriculum is that, even with the inclusion of all language skills, speaking and listening are not emphasized the way they should be, which is equally true of the examination process. While the Grade 12 national INJAZ examination should provide evidence of all the students’

successful efforts throughout the twelve scholastic years, the grades allocated for each part of this curriculum-oriented examination do not seem to support the claim (see Table 1). Moreover, the test items are aimed at eliciting information on students' reading and writing competencies exclusively, without any assessment of oral or aural skills.

Even more, the negligence of oral communication skills at the school leaving examination, which is administered by the MoE, lessens the motivation of both teachers and learners towards developing these skills. However, according to the National Palestinian Report (2016), during university admission procedures and on university courses, oral skills are included in tests. So, with the current situation, considerable inconsistency is found in language education, as English is only used in certain controlled and narrow contexts such as in-school classroom but not in exams or outside schools. However, some private schools and kindergartens have developed their own L2/L3 (English and/or French) multilingual curricula with the obvious objectives to develop students' language proficiency by relying on transfer as the dominant approach to teaching, while equal attention is paid to all skills in the different languages.

The majority of teachers speak Arabic as their first language, and the same is true for students, which is considered to be a yardstick in a more dynamic teaching approach. Currently, in the monolingual environment, learners receive little to no exposure to L2/L3 outside classrooms. In addition, the teachers set their goals based on standard native L2 varieties, which shrinks the learners' exposure to a wider range of varieties and dialects of English. This orientation is reinforced by resilience to change or to the adaptation of novel approaches to language teaching, or to making use of all the potential capacity available in a more modern and complex dynamic model of teaching. Confirming this perspective, Jenkins et al. (2011) state that “the notion that they should all endeavour to conform to the kinds of English which the native speaker minority use to communicate with each other is proving very resistant to change” (Jenkins et al., 2011: 308)

According to The Higher Education System in Palestine Report (2006), the enrolment and admission at all Palestinian higher education institutions follow more or less the same procedures. Among the minimum requirements for students to enrol in higher education institutions, the General Certificate of Secondary Education (INJAZ) or its equivalent for Palestinians from other countries is vital. Student placement in the faculties depends on the chosen stream indicated in the certificate. Therefore, INJAZ (General Certificate of Secondary Education, 12 years of schooling) is a high-stake examination.

High-stake examinations tend to have powerful ramifications including washback and test impact on educational processes and policymaking. Wall (1997)

distinguishes the two claiming that washback is “test effects on teacher and learner behaviour in the classroom whereas impact refers to wider test effects such as their influence on teaching materials and educational systems” (Wall, 1997: 100). When test results are used as a determining factor in the decision-making process in life, education and career, they are considered high-stake, which definitely bears high impact and subsequently washback effect on the behaviour of individuals and institutions alike.

Criteria of a good assessment instrument

In the framework of preparing a successful test instrument as a tool to measure achievements, development, or performance level, some characteristic features have to be maintained in order to claim the examination adequate. The major criteria that need to be met for creating a good test are its (1) validity, (2) reliability, (3) impact, (4) language task/test characteristics (5) where the first two criteria are considered to be the most important elements of an adequate assessment procedure (Gronlund, 1998; Gardner, 2012; Iseni, 2011). In the following, we will discuss the two criteria in detail; however, other features will also be mentioned in later sections of the study.

This study is concerned with the whole of the test instrument designed by MoE as a proficiency (or an achievement) test administered to school leavers, therefore, both test analysis and item analysis will be completed.

Test analysis is carried out through considering the criteria of *validity*, more precisely content validity, *reliability* and *impact*, while for item analysis, a framework of language task characteristics (Bachman & Palmer, 1996) is used to determine the *level of difficulty*, *discrimination power*, and the *quality of options* across the language skills included in the test. In the following part, we will elaborate on these criteria in more detail.

It is a widely held view to define *validity* as the test’s ability to measure what it is supposed to measure, however, this claim is expressed in different ways by different authors through time. Lado presents it as a matter of relevance as he wonders “Is the test relevant to what it claims to measure?” (Lado, 1961: 321), and he goes on to make recommendations on how to achieve maximum validity. Heaton states “a test is said to be valid if it measures what it is intended to measure” (Heaton, 1975: 153), which is later extended by adding “the extent to which it measures what it is supposed to measure and nothing else” (Heaton, 1988:159). Hughes extends the concept of validity as “an overall evaluative judgment of the degree to which evidence and theoretical rationales support the adequacy and appropriateness of interpretations and actions based on test scores” (Hughes, 1989: 20).

Another important feature of a good test is *reliability*. An instrument is said to be reliable if the same examination or test repeated after some time under the same

conditions and with the same participants yields highly similar scores and results for both sessions. Thus consistency, or stability, are key factors here, which means consistency from person to person, time to time or place to place. Joppe defines reliability as follows:

The extent to which results are consistent over time and an accurate representation of the total study population under study ... and if the results of a study can be reproduced under a similar methodology, then the research instrument is considered to be reliable (Joppe, 2000: 1).

Since, in this study, the GCSE (INJAZ) Examination is considered a high-stake exam, it is necessary to discuss its relation to education, so in the following part of this section, we will turn our attention to further principles which are also involved in the quality of a good examination in the wider context of the education process.

Some factors will affect the teaching process, and as a result, the test scoring process. Madaus (1988) defines a high-stake exam as a test whose results are seen – rightly or wrongly – by students, teachers, administrators, parents, or the general public, as being used to make important decisions that immediately and directly affect them, where phenomena such as *impact* and *washback* surely emerge. In the following, we will investigate the concepts of test washback, or backwash, and the test impact. However, since recently washback has been considered as a dimension of impact, we will see that the two concepts seem to overlap in a number of respects in literature.

Tests can influence what and how teachers teach and what and how learners learn in formal classroom settings. While washback is traditionally defined as the impact of tests on teaching and learning (Hughes, 2003; Green, 2015). Bachman (2004) extends the concept saying that washback can also be viewed as a subset of the impact of a test on society and on the educational system. This seems to be in line with the definition of impact as “any of the effects that a test may have on individuals, policies or practices, within the classroom, the school, the educational system or society as a whole” (Wall, 1997: 291).

The washback of a test “can either be positive or negative to the extent that it either promotes or impedes the accomplishment of educational goals held by learners and/or programme personnel” (Bailey, 1996: 268). Chen (2006) finds that some negative washback effects might result from the inefficiency of aligning the testing objective with a new curriculum, which would also hinder its implementation.

Based on professional literature, we can conclude that while the term washback, or backwash, is sometimes employed as a synonym of impact, most frequently it is used to refer to the effects of tests on teaching and learning at the classroom level.

As a means of evaluation, a test is often administered to get information about the student's improvement, or level, and to measure the result of the teaching-learning process. INJAZ test is, to a certain extent, an achievement test, which is held at the end of the teaching-learning process in one scholastic year, however, at the same time it demands and examines general language competencies developed throughout 12 years of learning English in a formal context, thus may be considered as a proficiency test as well. In other words, we may assume that this exam is a high-stake exam, which is intended to evaluate students' achievements after one year of learning English language and, at the same time, to evaluate students' overall achievements based on their accumulated knowledge from previous years in order to open the gates for pursuing undergraduate studies.

Being a high-stake examination, the INJAZ test instrument applied for measurement is required to meet further expectations of a good test considering the construction of items. Among the aspects of item analysis, we find that discussing their difficulty level, discrimination power, and the quality of options essential.

Among the criteria of evaluating item quality, the *difficulty level* of the item is concerned with how difficult or easy the item is for the test takers (Shohamy, 1985). Its importance emerges from the idea that simple and easy test items will provide us with little or no information about the differences within the test population. In other words, if the item is too easy, it means that most or all of the test takers provide the correct answer. In contrast, if the item is difficult, it means that most of the participants fail to provide the correct answer.

The *quality of options* also plays a key role in defining the difficulty of some of the test items. It emerges when the task provides the test takers with different options to choose from, for instance, when a task requires finding the correct answer in a multiple-choice question. According to Shohamy (1985), the quality of options is obtained by calculating the number of examinees who choose the alternatives A, B, C, or D or those who do not choose any alternatives. Accordingly, the test developers should be able to identify whether the distractors function appropriately or not so.

Another aspect of item analysis is the *discrimination power* that tells us about whether the item discriminates between the top group students and the bottom group students. Shohamy (1985) states that the discrimination index expresses the extent to which the item differentiates between top and bottom level students in a certain test.

Considering all the above criteria contributes to the effective evaluation of a test instrument, and provides feedback to develop the specific sections of the test items and the whole instrument, however, a detailed survey of item analysis seems to be beyond the scope of our article. Still some crucial aspects of item analysis will be considered in connection with our study area.

The purpose of the study

There have only been few studies that aim at discussing INJAZ English exam in Palestine with multiple aspects of the assessment procedures. Madbough (2011) measures the influence of the Tawjihi English exam on the Tawjihi students and teachers in Hebron concerning three domains: anxiety, output, and learning/ teaching methods. He recommends more training to be provided for teachers to enable them to modify their teaching methods and not to mix languages in teaching foreign languages, especially when teaching grammar. Also, he recommends that teachers should be offered more training on how to use tests in order to improve their instruction methods.

El-Araj (2013) aims to investigate the extent to which “Tawjihi” English language exam matches the standardized criteria of exams in Palestine between the years 2007-2011. She applies a descriptive-analytical approach employing two main techniques in her study, namely content analysis cards and a questionnaire, to elicit teachers' perceptions of Tawjihi English language exam. The results of the study show a low level of correlation between the exam contents and the textbook quality. Most of her recommendations emphasize the need for more ramifications and modifications of the exam as a holistic instrument developed for assessing all the language skills and for including higher-order thinking skills.

Amara (2003: 223) establishes some principles of the shift towards a more efficient Palestinian curriculum, more specifically, of the status of English language education in moving away from earlier conventions.

In modern language teaching and assessment, the *Common European Framework of Reference for Languages* (CEFR) has been established as a guideline for language professionals to enable them to provide standard education and teaching procedures internationally. According to Wall,

What [the CEFR] can do is to stand as a central point of reference, itself always open to amendment and further development, in an interactive international system of co-operating institutions...whose cumulative experience and expertise produce a solid structure of knowledge, understanding and practice shared by all (Trim, 2011: xi).

The CEFR provides a common basis for the elaboration of language syllabi, curriculum guidelines, examinations, textbooks, etc. across Europe. It comprehensively describes what language learners have to learn to do in order to use a language in communication and what knowledge and skills they have to develop so as to be able to act effectively. It also defines levels of proficiency which allow learners' progress to be measured at each stage of learning and on a life-long basis. Another purpose of CEFR levels is to assist self-assessment, so as language learners

can more clearly comprehend what they need to work on to reach the level they would like to achieve in their target language. The CEFR is designed to be applicable to many contexts, and it does not contain information specific to any single context. To use the CEFR in a meaningful way, test developers must also elaborate on the contents of the CEFR. Among others, this may also include considering which register of words or grammar might be expected at a particular proficiency level in a given language. It is agreed that the CEFR provides a shared base for developing the syllabi, exams, or textbooks, and determines levels of performance for measuring learners' progress or proficiency at different stages of acquiring the language. We claim that some principles of CEFR might be useful for our work in order to investigate the extent to which INJAZ test abides the common framework of foreign language assessment.

In our descriptive study, the 2018 science stream (module) version of the English Palestinian GCSE examination instrument is analyzed (1) in the light of the literature discussed above, and (2) in relation to the Common European Framework of Reference for Languages (CEFR) document in order to make conclusions on the qualities of the examination.

The study of INJAZ

According to Foorman (2009), tests are increasingly seen as means of evaluating school systems and measuring progress. Our experience suggests that communicative language teaching and testing principles are hardly maintained in the education processes in Palestine. Students who score high on written tests of English may find it difficult to communicate or express themselves orally.

According to The Palestinian Curriculum for General Education Report (1996), The INJAZ exam is assumed to (1) take all language skills (listening, speaking, reading, and writing) into consideration; (2) attempt to meet as many of the objectives for Teaching English as a Foreign Language (TEFL) assigned by the MoE as possible; (3) rely on the activities introduced in English for Palestine Grade 12; and finally, (4) consider the relative weight allocated for each skill in the teacher's book of English for Palestine when distributing the questions and the grades.

In the following, we will carry out the investigation of the test instrument of INJAZ 2018. First, we will discuss the general structure of the instrument, the structure of the papers included, skills involved, and the scoring system. Then we will move on to investigate the details of each paper, namely the test of reading skills, vocabulary, grammar, and writing skills. Finally, some statements on testing the oral communicative skills will be made.

Description of the examination structure

Communicative testing aims at focusing on real language use and learners' performance. Within this framework, language test items should include both objective questions and subjective questions to test students' creativity and their ability to express themselves freely (Foorman, 2009). Weir highlights that “integrative type of tasks, such as cloze tests, inform us about the student's linguistic competence but nothing directly about his/her performance ability” (Weir, 1990: 3).

GCSE English exams in Palestine focus on testing discrete-point language items which test a single item of the test-takers' language knowledge rather than integrative communicative competences. Table 1 presents the structural elements of the INJAZ English language exam and the assigned scores for each skill.

Table 1.
INJAZ (GCSE) 2018 English exam: format and scores distribution

Paper	Component	Question Number	Scores
Paper One	Reading Comprehension	Question Number One (20)	40
Paper Two	Reading Comprehension	Question Number Two (20)	
Paper Three	Vocabulary		25
Paper Four	Language	Question Number One (10) Question Number Two (10)	20
Paper Five	Writing		15
	Listening		Zero
	Speaking		Zero
Total			100

According to Table 1, listening and speaking skills are excluded. However, we believe that it is very important for aural and oral skills to be included in the exam, as English for Palestine is intended to be taught for communicative purposes. Also, students' competencies are supposed to be tested according to the principles of communicative language testing.

Concerning the structure of the test, it is also noted that the MoE's objectives are not fully met since the textbook activities do not seem to be in line with the exam tasks, and the distribution of scores is neither as clear as it should be nor as fairly balanced between objective and subjective questions as it should be. On the other hand, all the exam instructions are written in English as shown in Figure 1 below.



Figure 1: Exam Instructions

In order to use language tests and their scores to make inferences about language ability or to make decisions about individuals, Bachman & Palmer (1996) propose that test performance should correspond to non-test language use. This can be achieved by developing a framework of language use that enables teachers or test developers to consider the language used in Target Language Tests/Tasks (TLU) as a specific instance of language use, a test taker as a language user in the context of a language test, and a language test as a specific language use situation. Moreover, since the purpose of language testing is to enable teachers or test developers to make inferences about the test-takers' ability to use the language to perform tasks in a particular domain, Bachman & Palmer (1996) argue that the essential notions for the design, development, and use of language tests has to be analyzed in terms of the structure and form as well as the task specifications of the tests, as shown in Table 2 below.

Table 2. Test structure and test/task characteristics

Test structure	Comments
Number of parts/tasks:	The test is organized around 20 tasks within five papers, which contain 4 components—reading comprehension, vocabulary, language, and writing (see Table 1 above)—The purpose of these parts is to require the test takers to demonstrate their language competencies in all of these components. The test does not contain parts or components to test two skills, i.e. speaking and listening.
Salience of parts:	Parts are clearly distinct.
Sequence of parts:	It follows the sequence of previous years' tests (see: El-Araj, 2013; Madbouh, 2011).
Relative importance of parts or tasks	Each part independently stands out as important as any other part. There is no part that depends on the other.
Number of tasks per part:	Varied (5 tasks, 5 tasks, 5 tasks, 3 tasks, and 3 tasks distributed respectively).

Test task specifications	
Purpose	The purpose of the test is to arrive on students' overall achievements in the English language after 1 scholastic year which will be used as part of the students' local universities admission or receiving acceptance from universities overseas. Therefore, the test is considered a high-stakes test.
Setting	
Location, Materials and Equipment	Classrooms in schools. A pencil and an answer sheet
Time allotment	Two hours and thirty minutes (See Figure 1 above)
Language of instructions (LI)	The target language (i.e. English) because the test takers sit for an achievement test without any help from test invigilators.
Channel	Visual (writing) only
Instructions	The test-takers read the instructions silently (see Figure 1 & Appendix 1).

According to the test structure and task specification principles described in Table 2 above, the overall structure of the INJAZ exam is following a reproductive format rather than a communicative oriented one. Test takers have access to previous copies of the exams and may memorize the format and the strategies of the exam questions rather than preparing for an achievement, or proficiency test where their different language skills will be tested in an authentic manner. This yields poor results as students pay attention to the test form rather than following the exam instructions. According to the scoring method, the quality of options and difficulty level can promote fair assessment; however, the quality of options, when measured, may be negatively affected by the repetitive form of the test in general. The channel of the exam is visual i.e. written, which adds to the exclusion of the assessment for listening and speaking skills as illustrated in Table 1 above.

Testing Communicative Language Ability

One important dimension in language testing and assessment is the communicative competence of a language learner. Communicative language testing highlights the knowledge of the language and its application in real-life situations as the tasks are built upon communicative competence to involve other sub-competencies, represented in the grammatical knowledge of phonology, morphology, syntax as well

as socio-linguistic knowledge. Communicative competence is generally divided into three domains, which include (1) grammatical competence: words and rules; (2) pragmatic competence: both sociolinguistic and “illocutionary” competence, and (3) strategic competence: the appropriate use of communication strategies (Bachman, 1990; Canale & Swain, 1980). However, Canale & Swain (1980) establish a fourth competence type for cohesion and coherence of the text (i.e. discourse competence), which is not mentioned separately in Bachman (1990). As far as communicative language testing is concerned, Canale & Swain (1980) claim that the language user has to be tested not only on their knowledge of the language but also on their ability to use this knowledge in a communicative situation. In addition, Bachman & Palmer introduce two principles, which have become fundamental in language testing, namely (1) “the need for a correspondence between language test performance and language use” as well as (2) “a clear and explicit definition of the qualities of test usefulness” (Bachman & Palmer, 1996: 9).

The concept of language ability has traditionally been employed by language teachers and test developers alike, which incorporates the testing of four skills; listening, reading, speaking, and writing. These four skills have traditionally been distinguished in terms of channels (audio, visual) and modes (productive, receptive). Thus, listening and speaking involve the audio channel, which include receptive and productive modes respectively; while reading and writing are making use of the visual channel for the two modes (Bachman & Palmer, 1996). However, the authors argue that it is not sufficient to distinguish the four skills in terms of channel and mode only. Therefore, Bachman & Palmer (1996) propose to think in terms of specific activities or tasks in which language is used purposefully in replacement of ‘skills’. As a result, rather than attempting to define ‘speaking’ as an abstract skill, they believe it is more useful to identify a specific language use task that involves the activity of speaking, and describe it in terms of its task characteristics and the areas of the language ability it engages. Eventually, it can be more useful to conceptualize an oral communicative task as a combination of language abilities, including speaking, which lends itself to analysis through the description of task characteristics. This argument is presented in order to shed light on the exams’ items and their characteristics rather than on searching certain sections to spot certain skills.

Bachman & Palmer (1996) believe that their model of language ability can be used in the design and development of language tests. In order to facilitate this line of thoughts, Bachman & Palmer (1996) have found that it is useful to work with a checklist to help define the construct of an assessment one may want to complete in a given language test, i.e. the components of language ability rather than rigid language skills per se. Accordingly, this checklist can be used to judge the degree to

which the components of language ability are involved in a given test or test task. The sample of this study (GCSE English exam in Palestine) was analyzed according to (Bachman & Palmer, 1996) checklist as follows in Table 3.

Table 3. Components of language ability in GCSE English exam, Palestine

Components of language ability	Comments
GRAM: Vocabulary	general and technical (specifically related to Students' book)
GRAM: Morphology and syntax	Standard English
GRAM: Phonological/ Graphological	typewritten
TEXT: Cohesion	cohesive throughout
TEXT: Rhetorical organization	writing in essay form or statements
FUNCT: Ideational	Yes (Test takers express their ideational knowledge of English by describing, justifying, proposing, arguing, comparing, contrasting)
FUNCT: Manipulative	Yes (performing manipulative functions of making essays or statements)
FUNCT: Heuristic	Yes (working out solutions related to grammar and types of reading skill)
FUNCT: Imaginative	Not involved
SOCIO: Dialect	standard
SOCIO: Register	moderately formal
SOCIO: Naturalness	natural throughout
SOCIO: Cultural references and figurative language	events and procedures associated with Student Book and Work Book without figurative language

According to the checklist in Table (3) above, the GCSE English exam involves almost all the aspects of the checklist except having an imaginative function in the language characteristics of the test. Hence, the exam makers abide by the correct and scientific norms of preparing a high-stake test. In the following sections, the focus will be on describing each part of the test separately.

Testing Reading

Testing reading is categorized into two parts, namely, testing *reading accuracy* and testing *reading comprehension*.

In the accuracy test of the reading part, the focus is on the reader's ability to decode and utter written symbols of the language accurately while rendering chunks of sentences in a text using correct pronunciation properly. According to Zutell & Rasinski (1991), testing reading accuracy should follow a criterion that tests, on the

one hand, the reader ability to vary expression and volume that matches his/her interpretation of the text (expression). On the other hand, it generally tests smooth reading with some breaks, but word and structure difficulties are resolved quickly, usually through self-correction (smoothness). This skill is not included in the Palestinian GCSE English exam despite its importance for developing students' ability in pronunciation and speaking as a result.

Testing reading comprehension generally involves reading a text and asking students to answer inferential questions about the information implied in the text.

Other techniques including further skills, such as the ability to retell the story in the students' own words or to summarize the main idea or to elicit the moral lesson, might also be applied. Reading comprehension can be measured at three levels according to Karlin (1971), namely: testing literal comprehension, testing interpretive or referential comprehension, and testing critical reading. In testing literal comprehension, the focus is on testing skills such as skimming and scanning; meanwhile, testing interpretive or referential comprehension steers test takers to critically read and carefully analyze what they have read via comparing, contrasting or discussing. Lastly, testing critical reading, where ideas and information are evaluated critically, happens only if the students understand the ideas and information that the writer has presented e.g. testing students' ability to differentiate between facts and opinions.

On the first page of the Reading Comprehension section (see Appendix 1) of this exam, literal comprehension testing is demonstrated in the first two tasks (sub-questions 1&2) as well as referential comprehension testing can be identified in the rest of the tasks (3&4), however, no signs of other reading comprehension abilities, such as critical reading or thinking, can be found in the exam.

This section of the exam contains two types of questions, subjective and objective ones. The last task (number 5) is a referential comprehension testing type, however, the distribution of the scores is not clear for the section and the tasks. Moreover, the objective questions are assigned more scores than the more demanding subjective questions.

The second page (see Appendix 2) is also devoted to reading comprehension with 20 scores allotted to it. So, the total sum for *Reading Comprehension* is 40 scores (40% of the exam's total scores) with a random distribution of objective and subjective types of questions. Again, the unfair distribution of grades is obvious especially for the objective type of questions in comparison with language and grammar questions in the upcoming sections. This reading section contains reading a text from a source other than the coursebook, which is an advantage for testing students' capabilities to communicate meanings of vocabulary in unfamiliar and

external contexts. The instructions are only given in the English language as shown at the bottom of the page.

The items included in the reading comprehension section are aimed to require a mix of higher- and lower-level thinking skills, which render them to be of medium difficulty level. Although the same forms of questions are repeated over the years, the text itself is not extracted from the familiar textbook. The items range between different types of objective and subjective questions. In light of the scoring method employed, we can, therefore, assume that the paper is designed for both top group students and bottom group students, which may maintain good discrimination power. All exam items included in this section are compulsory and have to be answered.

Testing Vocabulary

Students' linguistic inventory is one of the yardsticks in language assessment studies. Harmer claims that "language structures make up the skeleton of a language while vocabulary is the flesh" (Harmer, 2003: 153). This claim implies that both language structures and vocabulary are equally important, however, interdependent. As a result, their assessment has been integrated into both progress tests and high-stake tests alike, especially in second-language teaching contexts.

Having good knowledge of English vocabulary is important for any language user. Therefore, 20% of the overall distribution of scores in the exam is allocated to this skill. The knowledge of vocabulary (See Appendix 3) is being tested within contexts that are familiar to students from the textbooks in the form of completion items, multiple-choice formats, and word formation items. However, students' efforts in completing vocabulary-related tasks and answering questions from the textbook do not require more sophisticated strategies, such as looking for contextual clues to decode the meaning of unknown words, noticing the grammatical function of the words, or learning the meaning of common stems and affixes, or other similar ones described by King & Stanley (2002), because of two reasons. Firstly, the task format only relies on the students' memorization skills, what is more, only includes objective questions. The other reason is that students are not taught the required strategies as part of their EFL journey in formal education.

In the phrasal verb question (question #C), students are asked to recall the memorized chunks of words to complete the matching task. In the following question (Question #D), all the words are provided together with distractors and the students are asked to choose the correct form to complete the sentences. The distribution of scores cannot be considered fair as the questions defined as higher-order thinking skills (HOTS) according to Bloom's taxonomy are assigned the same scores weight as the questions of lower-order thinking skills (LOTS). This is also true for the sample answer sheet provided by the MoE for the exam markers as their instructions

are not clear about how to score such questions and according to what principles are the scores added up or subtracted.

It is worth mentioning that the students make a lot of effort to memorize the English-Arabic meaning of words. However, these drills are not employed in the GCSE (Al-INJAZ) English exams. Instead, teachers and students alike use this technique to learn the meanings of words relying on their L1 in replacement of students' developing their individual vocabulary inventory.

For this section of the exam, the thinking skills required are of low level of difficulty as the items do not require higher-order thinking as much as the previous section does. In addition, all the vocabulary items presented in this section of the exam are directly related to the knowledge discussed in the textbook, and include objective type questions. Since dichotomous scoring method is applied, the index for the discrimination power in this part of the test would not be fair. As in the previous sections, all items included in this section are compulsory.

Testing Grammar

Grammar, named 'language skill' in this exam, is divided into two sections with 20% of the total exam scores (See Appendix 4). While Section A is fully compulsory, Section B allows students to choose two parts only. In Section A, Question A, and B are assigned one score for each correct answer. The scoring is fair in comparison with other skills' scores distribution, where students may have higher scores for the mere recollection of memorized items. However, many students may fail to answer some language questions correctly, because of misunderstanding the questions themselves. For example, in section B, where two tasks (out of three) are compulsory, even the text of the questions contains specialized words or structures above the students' expected language levels (e.g. reduced relative clauses or causative structure). A different and a simpler language could be employed to enhance students' understanding of this section also by providing an example for the intended task. During the test, the students receive written instructions both in their L1 (Arabic) and in English, nevertheless the instructions in Arabic appear only when a guideline is needed for choosing one or the other option on a list of topics.

The level of difficulty in this section is high as the items contain specialized concepts that students may not be familiar with or used in their mother tongue i.e. Arabic. Moreover, the items included in this section depend on the students' ability to employ the ideational function of the language such as comparing and contrasting. Besides, during the teaching process, students tend to memorize the grammatical forms according to some rules rather than drill them in communicative contexts, which hinders the students' ability to use grammatical forms creatively or to allow them to identify changes in these forms. This phenomenon affects their overall

achievements in this particular section of the test. The task items in this section vary between subjective and objective types which are all compulsory. As this section includes items requiring both higher- and lower-level thinking skills it addresses both bottom and top-level test-takers.

Testing Writing Skill

Writing is considered to be the most advanced skill, which is based upon proficiency in a number of specific knowledge areas and sub-skills such as grammar, sentence order, vocabulary, spelling, etc.

Weigle (2002) makes a distinction between two forms of testing writing performance, which are the holistic and the analytical one. She explains that “in analytic writing, scripts are rated on several aspects of writing or criteria rather than given a single score. Therefore, writing samples may be rated on such features as content, organization, cohesion, register, vocabulary, grammar, or mechanics” (Weigle, 2002: 114). However, here she also claims that “on a holistic scale, by way of contrast, a single mark is assigned to the entire written texts” (Weigle, 2002: 114) and concluding by claiming that a holistic scale is less reliable than an analytic one.

In order to test writing skills, Aryadoust (2004) presents the criteria (see Appendix 6) which we follow to assess the writing tasks in this exam.

In the writing section, two guided writing tasks are provided, and the students need to choose only one of them (see Appendix 5). Traditionally, students are accustomed to these types of questions available in previous GCSE English exams where they memorize the forms of the questions and the answers. Hence, writing does not emerge as a fully productive skill but rather as a mere repetition and a blind imitation of previously prepared and memorized forms that fit in for any content of the writing section. This may lead to learners practising and preparing merely to pass the exam. This attitude may result in unanticipated, harmful consequences of a test (a negative washback effect) since the students may focus too heavily on test preparation at the expense of actually developing an independent capacity to produce their thoughts in a written form.

The items included in the writing section require higher-order thinking skills. The problematic issue of this section is that students develop a sense of dependency as they memorize the expected writing compositions from the textbook and do not bother to prepare for an academic writing process that truly reflects their level of competences in writing. Students tend to employ some previously prepared clichés that may fit in for such type of questions. Therefore, the aim of the test is not met as students depend on memorizing rather than on developing their own ideas.

On the other hand, since the two questions are of subjective type, they may hinder low achievers from getting high scores in this section as their language capacity will

probably not facilitate efficient writing. Due to this fact and the scoring method backed up by high faculty values, the two questions are considered to demonstrate a high difficulty level.

Testing Aural and Oral Skills (Speaking and Listening)

Lundsteen (1979) argues that listening skill, similarly to reading comprehension, is commonly defined as a receptive skill comprising both a physical process and an interpretative, analytical process. However, the definition of ‘analytical process’ incurs some critical listening skills, such as analysis, and synthesis in addition to evaluation. Nonverbal listening, which includes comprehending the meaning of the tone of voice, facial expressions, gestures, and other nonverbal cues belong to the same umbrella term, as well. No one can argue against the importance of listening skill in communication or of the assessment of it in language education. And the same is valid for speaking.

Mead & Rubin (1989) describe speaking skill according to the following aspects: (1) communication activities that reflect a variety of settings: one-to-many, small group, one-to-one, and mass media; (2) using communication to achieve specific purposes: to inform, to persuade, and to solve problems, and finally, (3) basic competencies needed for everyday life: giving directions, asking for information, or providing basic information in an emergency.

Given the importance of oral communication, we can claim, unfortunately, both listening and speaking skills are excluded from the testing process of the Palestinian GCSE English exam.

Conclusion

In our paper, we claim that GCSE (INJAZ) 2018 English Exam seems to fail to meet the criterion of validity for several reasons.

The individual items and parts in the GCSE (INJAZ) 2018 English Exam do not meet all the standards and objectives set by the Ministry of Education. This means, the exam cannot be considered to be valid in its content, as many questions and tasks – initially designed to test students’ general language abilities – fail to examine all language skills or aspects. Instead, they often test students’ ability to memorize and recall various types of language information and the ability to use some techniques. This partly support Sun's (2000) results, which suggest that some items of the test should be improved to match the course students are taught.

Ramahi (2018) stresses that the formal education in Palestine is not adequately responding to the challenges and demands of the political and socioeconomic conditions, neither at the curriculum content level nor the modes of assessment.

This is due to different factors, first of all to the teaching methods applied, which are often exam-oriented, and more particularly to factors revealed in this study, that is to the fact that the exam papers themselves do not seem to employ any aural or oral skills questions. Moreover, GCSE English exams in Palestine focus on testing discrete-point language items which test a single item of the test-takers' language knowledge rather than integrative communicative competences.

Among the main reasons for learning a language, a dominant element is to practise it in daily communication mainly through listening and speaking. GCSE (INJAZ) 2018 English Exam fails to focus on different types of competencies mentioned in (Canale & Swain, 1980; Bachman, 1990; and Bachman & Palmer, 1996) as two of the important language skills required to assess these competencies are neglected in the exam. This could be avoided in the upcoming years with fair attention given to cover all the skills without compromising both the speaking and listening skills and better modification to the scoring method. In other words, the exam lacks validity of content as it does not measure what it is supposed to measure, i.e. listening and speaking skills in authentic situations as stressed by the CEFR framework.

Despite the repetitive forms of the exam, it does not guarantee that the students will get the same or higher scores when retaking the test. Repetition may help test-takers pass the test but fails to provide information concerning their real communicative competencies.

Our experience suggests that the exam has exerted a negative washback that affects the teaching and learning processes in a negative way. The education seems to have shifted to prepare students to pass the exam rather than to improve their language competencies. During preparation, students do not focus on the different parts of the curriculum but rather on some tactics that may assist them to complete the exam tasks more easily. A further implication of the washback effect observed may result in reconsidering the curriculum itself. The question of whether the content therein needs redesigning or revising based on recent linguistic research and development, or in harmony with the changes in the social and economic conditions in the environment, is still open.

Our current findings seem to be in line with the results of previous studies (Madbouh, 2011; El-Araj, 2013) in assessing GSCE exam in Palestine.

In summary, we have found that the INJAZ English Exam fails to meet the content validity criterion of a good test, and in addition, it seems to have affected the teaching and learning processes negatively with steering the teaching and learning procedures towards the aim of students' passing the exam rather than towards teaching and learning the English language.

More precisely, each section of the test fails to meet the requirements to render a good test. The reading section lacks a fair distribution of tasks and scores for critical

reading or thinking. The vocabulary section only includes objective questions requiring highly sophisticated strategies. The instructions of some tasks in the grammar section are beyond the students' expected level of competences, thus should be more finely graded so as students understand what they are exactly required to do in each task. The question format in the writing section is repeated over the years, which is resulting in a negative washback effect. There are no sections to assess the students' performance in oral communication tasks, in other words to assess their competencies in speaking or listening.

As a result of the above-mentioned deficiencies, the instrument fails to maintain the standards of CEFR and the criteria of a good test, which are vital elements of the MoE TEFL objectives, as well. It would be essential for language test developers "to probe more deeply into the nature of the abilities we want to measure" (Bachman, 1990: 297), and to provide a critical evaluation of the validity of the measurement tools (Shohamy, 1998).

In conclusion, a systematic revision of the test components should be considered in light of the MoE TEFL objectives to avoid the above-mentioned deficiencies of the test in the upcoming years.

References

- Allen, D.** (1998) *Assessing Student Learning*. New York: Teachers College Press.
- Amara, M.** (2003) Recent Foreign Language Education policies in Palestine. *Language Problems & Language Planning* 27/3. 217-232.
- Aryadoust, V.** (2004) Investigating Writing Sub-skills in Testing English as a Foreign Language: A Structural Equation Modeling Study. *The Electronic Journal for English as a Second Language*. 1-20.
- Bachman, L.** (1990) *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.
- Bachman, L.** (2004) *Statistical Analyses for Language Assessment*. Cambridge: Cambridge University Press.
- Bachman, L., & Palmer, A.** (1996) *Language Testing in Practice*. Oxford: Oxford University Press.
- Bailey, K.** (1996). Working for Washback: a Review of the Washback Concept in Language Testing. *Language Testing*, 13/3. 257–279.
- Bernard, J., & Maître, J.** (2006) *Studies on the Palestinian Curriculum and Textbooks: Consolidated Report*. UNESDOC.
- Canale, M., & Swain, M.** (1980) Theoretical bases of communicative approaches to second language teaching and testing. *Journal of Applied Linguistics*, 1/1. 1-47.
- Palestinian Curriculum Center** (1996) *The First Palestinian Curriculum for General Education (in Arabic)*. Ramallah: PCDC.
- Chen, L.** (2006) Washback Effects on Curriculum Innovation. *Academic Exchange Quarterly*. 206-210.
- The European Commission** (2006) *The Higher Education System in Palestine, National Report*. The European Commission. Retrieved from http://www.reconow.eu/files/fileusers/5140_National-Report-Palestine-RecoNOW.pdf
- El-Araj, T.** (2013) *Evaluating Twjehi English Language Exam in Palestine over the Last Five Years*. Unpublished Dissertation.

- Foorman, B.** (2009) The Timing of Early Reading Assessment in Kindergarten. *Learning Disability Quarterly*, 32. 217-227.
- Gardner, J.** (2012) *Assessment and Learning*. London: SAGE Publications Ltd.
- Green, A.** (2015) The ABCs of Assessment. In: Tsagari, D., Vogt, K., Froehlich, V., Csépes, I., Fekete, A., Green, A., . . . Kordia, S. *Handbook of Assessment for Language Teachers*. TALE project. 1-15
- Gronlund, N.** (1998) *Assessment of student achievement*. Boston: Allyn & Bacon.
- Harmer, J.** (2003) *The practice of English language teaching*. Essex, England: Longman.
- Heaton, J.** (1975) *Writing English language tests*. London: Longman Group UK Ltd.
- Heaton, J.** (1988) *Writing English language tests: Longman Handbook for Language Teachers (New Edition)*. London: Longman Group UK Ltd.
- Hughes, A.** (1989) *Testing For Language Teachers*. Cambridge: Cambridge University Press.
- Hughes, A.** (2003) *Testing Language for Teachers*. Cambridge: Cambridge University Press.
- Iseni, A.** (2011) Assessment, Testing and Correcting Students' Errors and Mistakes. *Language Testing in Asia*, 1/3. 60-90.
- Jenkins, J., Cogo, A., & Dewey, M.** (2011) Review of developments in research into English as a lingua franca. *Language Teaching*, 44. 281–315.
- Joppe, M.** (2000) *The Research Process*. Retrieved February 25, 1998, from <http://www.ryerson.ca/~mjoppe/rp.htm>.
- Karlin, R.** (1971) *Teaching Elementary Reading: Principles and Strategies*. Harcourt Brace and Jovanovich, Inc.
- King, C., & Stanley, N.** (2002) *Building Skill for the TOEFL*. Edinburg: Thomas Nelson Ltd.
- Lado, R.** (1961) *Language Testing: The Construction and Use of Foreign Language Tests: a Teacher's Book*. Bristol, Inglaterra: Longmans, Green and Company.
- Lundsteen, S.** (1979) *Listening: Its Impact On Reading And The Other Language Arts*. Washington, DC: National Council of Teachers of English, Urbana, IL.
- Madaus, G. F.** (1988) The Distortion of Teaching and Testing: High-Stakes Testing and Instruction. *Peabody Journal of Education*, 65/3. 29-46.
- Madbouh, A.** (2011) *The Tawjihi English Exam Washback on Tawjihi Students and Teachers in Hebron, Palestine*. Unpublished dissertation.
- Mead, N. A., & Rubin, D. L.** (1989) *Assessing Listening and Speaking Skills*. ERIC Digest.
- National Palestinian Report** (2016) *The Higher Education system in Palestine, National Report*. The European Commission.
- Nicolai, S.** (2007) *Fragmented Foundations: Education and Chronic Crisis in the Occupied Palestinian Territory*. UNESCO: International Institute for Educational Planning.
- Ramahi, H.** (2018) *Education in Palestine: Current Challenges and Emancipatory Alternatives*. Ramallah, Palestine: RLS Regional Office Palestine & Jordan.
- Shohamy, E.** (1985) *A Practical Handbook in Language Testing For the Second Language Teacher*. Tel Aviv: Tel Aviv University.
- Shohamy, E.** (1998) Critical language testing and beyond. *Studies in Educational Evaluation*, 24/4. 331-345
- Sun, C.** (2000) Modern language testing and test analysis. *Journal of PLA, University of Foreign Languages* 23/4. 82–86.
- Trim, J. L.** (2011) Some Earlier Developments in the Description of Levels of Language Proficiency, preface to Green. In: Green, A., *Language Functions Revisited: Theoretical and Empirical Bases for Language Construct Definition Across the Ability Range*. Cambridge: Cambridge University Press. xxi-xxxiv
- Wall, D.** (1997) Impact and washback in language testing. In: Shirai, Y. (auth.), G. R. Tucker, & D. Corson (Eds.), *Encyclopedia of language and education*. Dordrecht: Kluwer Academic Publications. 291–302.
- Weigle, S.** (2002) *Assessing writing*. Cambridge: Cambridge University Press.
- Weir, C.** (1990) *Communicative Language Testing*. Englewood Cliffs: Prentice Hall.

- Yamchi, N.** (2006) English Teaching and Training Issues in Palestine. *TESOL Quarterly*, 40/4. 861-865.
- Zutell, J., & Rasinski, T.** (1991) Training teachers to attend to their students' oral reading fluency. *Theory Into Practice*, 30. 211-217.

We acknowledge the financial support of Széchenyi 2020 under the **EFOP-3.6.1-16-2016-00015**.

Appendix 1 Reading Comprehension

English language
Paper: ---
Time: 2:30 Hours
Total Marks: (100)

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ
امتحان الشهادة الثانوية العامة – الإجازة
G.S.C.E 2018
Science and
Entrepreneurship
AI-INJAZ


دولة فلسطين
وزارة التربية والتعليم
الإدارة العامة للقياس والتقويم والامتحانات
Date: 4/6/2018

Reading Comprehension (45 points)

Question Number One (25 points):

Read the texts and do the tasks below.

_____ There are limits to how many sports can be included in the Olympic Games. In the 2012 Summer Games, there were 26, and the International Olympic Committee (IOC) have added two more for 2016 (golf and rugby sevens). There are various criteria that the IOC bases its decisions on, but the most important is probably international popularity. Sports tend to get included if **they** have a tradition and are played and watched by enough people in enough countries.

_____ This leads to some interesting questions. Why, for example, was baseball included from 1984, but then dropped for 2012? Although some think it is only popular in the USA, it actually has a large following in South America and Asia too. The main reason for leaving it out, according to some, is a combination of geography and politics: there is little interest in **the sport** in Europe, and European members are in a majority on the IOC.

_____ Of the other sports not (yet) included, one that stands out is squash. It has a long tradition and is played all over the world. It's also more obviously a real sport than some other Olympic 'sports'. In fact, it's more mysterious why certain sports are included. Take speed walking, for example, a long-standing Olympic regular. The rules say that the athletes have to have both feet in contact with the ground at all times, and must keep their legs straight. The result, to me at least, looks very silly.

_____ More importantly, it breaks one of my two personal rules for judging serious sports because it puts artificial limits on what can be done. My other rule is that the scoring has to be objective and measurable, not a matter of opinion.

1. Match the headings with the suitable paragraph above:

(6 pts)

1. Two examples of strange decisions
2. What makes an Olympic sport?
3. My own criteria
4. A change of mind

2. Decide whether each of the following is True or False:

(4.5 pts)

1. There are 26 sports in the 2016 summer games. ()
2. Baseball is well-known in most European countries. ()
3. The writer suggests that speed walking should not be included in the Olympics. ()

3. Complete the following:

(5 pts)

1. The IOC included and For the 2016 Olympic games.
2. For the IOC the most important reason for including sport is
3. Baseball was left out from 2012 Olympic games for and reasons.

4. Answer the questions:

(5 pts)

1. According to the writer, why should squash be considered as a game in the Olympics?
A B C.....
2. What are the writer's criteria for scoring in sports?

5. Write what the following words and pronouns might refer to:

(4.5 pts)

- a. **they**
2. **the sport**
3. **their**

Appendix 2 Reading Comprehension (Paper two)

Two: Read the text and find answers to the questions below.

(20 pts)

A famous Chinese proverb says: ‘Pearls don’t lie on the seashore. If you want one, you must dive for it by limiting yourself to what you already know, you are likely missing out on professional opportunities, life experiences, and personal growth. To become in favor of taking these chances, learn to think positively about stepping out of your comfort zone.

Tip 1: In other words, what are the things that you believe are worth doing but are afraid of doing yourself because of the potential for failure. Draw a circle and write **these things** down outside the circle. This process will keep you away from disappointment. Moreover, it will allow you to clearly identify, not only your discomforts but also your comforts. Write identified comforts inside the circle.

Tip 2: There is no substitute for this step. If you want to become better at something, you must start hanging out with the people who are doing what you want to do and start imitating **them**. Almost inevitably, their influence will be of great importance on your behaviour.

Tip 3: Don’t try to jump outside your comfort zone, you will likely become negatively affected and jump right back in. Take small steps forward the fear you are trying to overcome. If you want to do public speaking, start by taking every opportunity to speak to small groups of people. You can also practice with family and friends.

Tip 4: many of us are so afraid of failure, that we would rather do nothing than take a shot at our dreams. Begin to treat failure as a teacher. First, you should learn from experience. Then you take that lesson to your next adventure your chance of success.

1. The main idea of paragraph one is:

(2 pts)

- a. take baby steps
- b. hang out with risk takers
- c. become aware of what’s outside your comfort zone.
- d. see failure as a teacher

2. Decide whether the following is True or False

(3 pts)

- 1. Identifying the possible discomforts is one of the necessary tips to avoid the feeling of disappointment. ()
- 2. You should stay away from risk takers because they will affect you badly. ()
- 3. Fear of failure will inevitably increase our chances of success in life. ()

3. Complete the following

(5 pts)

- 1. Staying in your comfort zone, you are likely to miss out on, and
- 2. The writer’s opinion in Tip 3 is to avoid stepping outside your comfort zone because you’ll probably and

4. Write suitable tips from the text to the people in the following situations:

(6 pts)

Situation	Suggested tips
I want to want to do public speaking but I’m afraid I cant.	a. b.
How can I turn failure into big success?	a. b.

5. Choose the correct answer.

(2 pts)

- 1. Pearls don’t lie on the seashore. If you want one, you must **dive** for it. The word **dive** means:
 - a. take risks
 - b. think
- 2. To get out of our comfort zone, the writer suggests that we should
 - a. hump out quickly
 - b. step out gradually

Appendix 3 Vocabulary

6. Write what the following words and pronouns refer to: (2 pts)

1. those things 2. them

Vocabulary (20 points)

A. Replace underlined parts of the sentences below with words from the text. (5 pts)

point	reveal	clinging to	straight	selfish
-------	--------	-------------	----------	---------

- I felt tired when I got home, so I had gone immediately to bed.
- I'm not going to this afternoon's session. I don't a purpose of it.
- She never considers anyone but herself. She's totally unaware of the people's feelings.
- The rescue team discovered four men holding on tightly a piece of wood in the sea.
- Doctors are nor allowed to show confidential information.

B. Complete the sentences with words from the box. (5 pts)

humanitarian	tragedy	impressed	excuse	capable
--------------	---------	-----------	--------	---------

- I've never known him to miss a meeting. I'm sure he will have an
- The pilot avoided a when he succeeded in preventing the plane from crashing.
- The film director was so by her performance that he directly offered her a job.
- The employee isn't well qualified, so he is not of doing such kind of work.
- The United Nations is sending a aid to the areas worse affected by the war.

C. Match verbs with particles from the box to make phrasal verbs, then use them to complete the sentences. (4 pts)

choose	out
work	between
throw	across
came	away

- Try to all what you don't need. You have a pile of unnecessary things.
- I a vase exactly like yours in an antique shop.
- If you all the costs of the project well, I think you'll achieve a great success.
- She had to giving up her job or looking after her family.

D. Choose the correct answer. (6 pts)

- The company is (**on / in**) danger pf having to close.
- The company has a large (**marketing/market research**) department that designs advertisements to help customers to choose the suitable brand.
- Most people prefer governmental field to work so as to get job (**security / description**)
- Sorry, we were late because we had (**break down/break through**) in the car while driving to work.
- Dr Samuel Holliday was one of the (**cofounders / coauthors**) who wrote the report.
- There has been a (**cutback/ breakdown**) in the government spending on new projects.

Appendix 4 Grammar

Language (20 points)

A. Complete the sentences with the correct tense of the verbs in brackets. (4 pts)

1. I (not agree) with what he usually (say)
Why are you holding a piece of paper?
2. I (write) a letter to my friend back home in Palestine.
3. She (have) a beautiful dream when the alarm clock (ring)
4. Why you (look) at me like that? Have I said something wrong?
5. Tamer (read) a book for two hours and he (read) pages so far.

B. Rewrite the following sentences using the words between brackets. (4 pts)

1. Why didn't you take these tablets before leaving? They asked
2. Is your school far from here? Someone asked
3. Don't talk on the mobile while driving. The police warned me against
4. Mary didn't prepare for the interview, so she didn't get the job. Mary wishes

C. Circle the correct answer. (2 pts)

1. Muneer was doing his work and then he stopped (**to watch / watching**) the football match on TV.
2. Our neighbors apologized for (**making / to make**) a loud noise.
3. He (**must have been / must be**) out because he isn't answering the phone.
4. I think we're lost. We (**should have brought / might have brought**) a map with us.

From this section answer two parts only.

1. Complete the sentences with {a / an / the / - no article} (5 pts)

1. We went back to area where I was brought up and got working job for the government.
2. The secretary lost important document and was fired.
3. She has been studying business law for two years.
4. Have you ever gone hiking in Alps?
5. Brazil is the largest country in South America.

2. A. Rewrite the sentences using reduced relative clauses. (2 pts)

1. I come from a city which is located in the southern part of the country.
2. Students who hand on their essays late will be punished.

B. Rewrite the sentences, replacing the underlined part with a causative structure (have + object + past participle). (3 pts)

1. We'd better ask someone to check the letter before you send it.
2. Don't forget to take your jacket to the cleaners before the wedding.

3. Correct the sentences. (5 pts)

1. They'd replace the offer soon, don't they?
2. His explanation to the problem isn't convinced.
3. I object to say sorry for something I didn't do.
4. Bill used to being fit. Now he's in a terrible condition.
5. I like everything apart of getting up so early.

Appendix 5 Writing

Writing (15 points)

Choose one of the following topics.

اكتب موضوعا واحدا من أحد الموضوعين

1. **Older generation find it quite exciting to live in another country but of course there are things that they'd miss, whereas most teenagers seem to get used to changes quite quickly. Write a short essay about the above statement.**

- A. Write about positive points of living in another country.
- B. Write about negative points of living in another country.
- C. State your opinion and say why they have these feelings.

Start your email with the following address:

2. **Time management is the shortest route to success.** To what extent do you agree?

- A. Why do we need to manage our time?
- B. How can we manage our time properly?
- C. What benefits could we get from time management?

Appendix 6 (Aryadoust, 2004) Criterion and Descriptors to Assess and Score Writing Samples

Criterion (sub-skill) Description and elements	Criterion (sub-skill) Description and elements
Arrangement of Ideas and Examples (AIE)	presentation of ideas, opinions, and information. aspects of accurate and effective paragraphing. elaborateness of details. use of different and complex ideas and efficient arrangement. keeping the focus on the main theme of the prompt. understanding the tone and genre of the prompt. demonstration of cultural competence.
Communicative Quality (CQ) or Coherence and Cohesion (CC)	range, accuracy, and appropriacy of coherence-makers (transitional words and/or phrases). using logical pronouns and conjunctions to connect ideas and/or sentences. logical sequencing of ideas by the use of transitional words. the strength of conceptual and referential linkage of sentences/ideas.
Sentence Structure Vocabulary (SSV)	using appropriate, topic-related and correct vocabulary (adjectives, nouns, verbs, prepositions, articles, etc.), idioms, expressions, and collocations. correct spelling, punctuation, and capitalization (the density and communicative effect of errors in spelling and the density and communicative effect of errors in word-formation. appropriate and correct syntax (accurate use of the verb tenses and independent and subordinate clauses) avoiding the use of sentence fragments and fused sentences. appropriate and accurate use of synonyms and antonyms.