

---

## **Evaluating 2020-2021 and 2021-2022 GCSE English Exams and Exam Method against Related Curriculum Goals and Stakeholders' and Experts' Perceptions**

**Emad Albaaly (Ph.D.).**

Associate Professor of Curriculum and Instruction of English, Faculty of Education,  
Suez Canal University, Ismailia, Egypt

ORCID <https://orcid.org/0000-0002-7003-8747>

---

**Citation:** Albaaly E. (2022) Evaluating 2020-2021 and 2021-2022 GCSE English Exams and Exam Method against Related Curriculum Goals and Stakeholders' and Experts' Perceptions, *International Journal of English Language and Linguistics Research*, Vol.10, No 4, pp.33-75

---

**ABSTRACT:** *Evaluation instigates learning and teaching and is central to them. All elements in the educational process have to be evaluated, including the evaluation of exams and related exam methods. This study aimed at evaluating the Egyptian GCSE English exam and exam method against GCSE curriculum goals and stakeholders' (i.e. teachers' and students') perceptions of the exams. The study adopted a qualitative approach: a fixed alternative questionnaire and a semi-structured E-interview both administered to participant experts and practitioner teachers (n 50 and 5, respectively), and it also incorporated designing and using another questionnaire (in two versions) administered to GCSE graduates: 50 for the school year 2020-2021 and 100 for the 2021-2022. Results revealed that certain curriculum language skills and elements were absent, and the vast majority of them were under-represented and improperly addressed. Even, the curriculum novel content was covered inappropriately. It was then found out that the exams had few strengths and many weaknesses, regarding the question type, language skill and language element representation/coverage, exam nature and appropriateness for students, incompliance with curriculum and OBD. Other exam implications were covered. Plans and recommendations for intervention included inclusion of performance-based questions for three language skills and the translation element, coverage of grammar and vocabulary, inclusion of language functions and idioms, variation of question types appropriate for language skill/element, and commitment to the OBE approach in the exams as well as to research.*

**KEYWORDS:** exam evaluation, assessment method, Egyptian exams, English Exam, Ministry of Education, GCSE, GCSE graduates, Secondary Stage

---

## **INTRODUCTION**

The Ministry of Education (henceforward referred to as M.o.E.) adopted a new method for assessing GCSE Students' levels in all subject areas from the year 2021. They presume the new method helps assess the students' understanding, application, and analysis qualities (Oxford Business Group, 2018). They sought a development and employment of learners' skills rather than pure memorization and call-up of knowledge (OECD, 2015: 13). Apparently, they also wanted the classroom to be rather interactive, and they aspire after a change in student assessment method which demonstrate students' skills and equip those students with pre-tertiary requirements (Ibid, 2015: 13-

14). It is also well known that curriculum attainment of goals and relating objectives are based on effectuating, addressing, and evaluating methods of assessments (Islam et al., 2021; Ewiss, 2021).

After two years of the administration of the new M.o.E. assessment method at the end of the Secondary Stage, there emerged a need for it to be evaluated, as evaluation is pivotal to the education process (Robiashi and Lestari, 2020: 80). As it is known, evaluation of assessment methods adds some supporting remarks or recommendations for improvement (Srinivasan, 2016:1). To the study, evaluation cares about effectuating assessment findings and turning them a reality. Of course, evaluation is different from *measurement* which, in educational terms, means quantifying events so as to give a quantitative meaning to the qualities imbedded. Evaluation is different from *assessment*, too. As assessment describes the progress students have achieved towards attaining a certain goal at a certain time and does not explain why something happened or how actions could be improved and it does not make recommendations as does evaluation (Ibid:1-2).

Thus, evaluation of assessment methods is inevitable to the general quality in education (Ewiss, 2021: 7). In order to recognize how a system works, it is imperative that the opinions of the clients/stakeholders who either receive or address the service be taken into consideration. It is widely known that the success of school performance must be based on the criteria of evaluation (Ewiss, Ibid: 7). Spady (1994: 21) states four principles for good assessment, or what is called 'outcome-based education' (OBE). The first is connected with having a clear objective regarding the necessity of assessing students' actual abilities. The second principle is to enable students to identify what they need aims to attain. The third one informs that students' activities must be engaging and hard so that higher thinking is involved. Students' must be evaluated on all types of performed tasks and practical aspects, using portfolio, assignments, and all types of ongoing assessment. The fourth procedure includes a belief in students' individual different abilities and character. Students differ in pace and abilities. Therefore, they should be tested at a time when their optimal abilities have been reached, and test time should be re-occurring to fit this process.

Generally, relating to the third point/principle above, when a curriculum is focused on gaining and understanding knowledge and when it focused on critical thinking and problem-solving skills (i.e. higher-thinking), students' assessment must reflect these processes imbedded in the curriculum (Krafft, 2012; Latha and Ravichand, 2019).

Thus, in order to understand the nature of assessment of the General Certificate for Secondary Education (GCSE)/ third-year Secondary Stage students for evaluation purposes, their curriculum components (i.e. subjects taught), must be understood first.

Obviously, the GCSE students are divided into three domains: literary, scientific, and mathematic. All domains study *compulsory across-domain* subjects which include religious education, Arabic and foreign languages, as well as statistics, national education, and one elective subject out of three: agriculture, industry, and home economics. There are also specialised *compulsory by domain* subjects, e.g. chemistry, biology, geology, environmental sciences (for the scientific domain); geography,

philosophy, psychology, logic, science, and sociology (for the literary domain), and algebra and solid geometry, differentiation, statics and dynamics (for the mathematic domain). It is worth mentioning that the *compulsory by domain* subjects and all languages from the *across domain subjects* are part of the students' whole score (%100). The rest of *across domain subjects*, thought unincluded in the total score, are just compulsory for students to pass.

Related to this classification of subjects is Ewiss' (2021: 9). However, unlike the above-mentioned one, Ewiss (Ibid:9) states that "GCSE students study optional specialised subjects: Students study chemistry, biology, geology, environmental sciences, geography, science, philosophy, logic, psychology, and sociology." In fact, however, it is seen that these subjects are not '*optional*' subjects; they are, actually, as stated earlier, '*compulsory by domain*' ones. This has been taken for granted in Egypt even before the new Exam method is adopted.

As for the grounds of change, it is thought that what led to the change in the evaluation system/ exam method at the M.o.E. was the assumption that no analysis or criticism of the curriculum was allowed and the curriculum itself contained much about history rather than advances in science and technology and relating methods of thinking, which caused problems to the stakeholders of the educational process as well as the society (Oxford Business Group, 2018; Saavedra, 2019). Apparently, the ministry wanted a shift from memorization-based student answers/ exams to higher thinking. Ewiss (2021:11) believes that the previous evaluation process, i.e. prior to the new method introduction, merely depended on paper and pen tests and measured merely knowledge and the least number of thinking skills. It is believed that these above-mentioned reasons formed solid grounds for a change and the ministry stated the new method was into effect from the beginning of the year 20201 and students would do the learning and the assessment, accordingly. The new method introduced E-exams, which were later abandoned, and only multiple-choice questions for assessing students' learning at the end of the year.

### Challenges to change

Essentially, from an administrative point of view, it is known that whenever there is a new assessment method which is about to be adopted, teachers and school administrators must receive training on the new method in order to optimize the results. However, this is claimed not to have taken place. Evidence from research supports this claim. For example, Ahmed (2019) finds out that teachers needed training to gain necessary about and be skillful at the educational technology to be used in the assessment projects to be adopted, such as tablet devices, for example. The claim is also supported by Hasan et al (2021) which finds that administrators, too, lacked the knowledge and skills relating to leadership and planning, technical supervision, electronic administration, teacher professional development, and evaluation system.

O'Loughlin (2006: 72) claims that the main *role* of assessment is originally to "bring progress in education". Clearly, all program-based education depends on assessment results. A superior educational program depends on the assessment results of an inferior

one. Robiashi and Lestari (2020: 80) support O'Loughlin's views and they add that assessment is considered an integral part in the education process.

Obviously, there is a distinction between two particular *types* of assessment to be mentioned in this concern. Formative assessment, as a type, means assessing students periodically to ensure progress in learning (Ibid: 81). The other type called summative assessment, stands for end product final assessment of students' levels (Nurwahidah et al, 2022; 185). The GCSE exam is an example of this type. Summative results (Ibid) can show more reliable data regarding the final outcomes than formative findings. Summative assessment is seen by others as means to check whether students have satisfied the learning objectives or attained the competence level required, e.g. Abduh (2021).

#### **Multiple-choice assessment method**

It is well known that for the national exams of the years 2021 and 2022, the M.o.E. adopted using the sole type of multiple-choice questions for assessing not only the GCSE English subject but also the majority, if not all, of the other subjects mentioned earlier as well, to the best of the researcher's knowledge.

In the following section, a research-based examination on whether the sole-type multiple-choice assessment method for the English subject, being the particular focus of this study, and the deliberate negligence of the assessment of listening, actual act/performance of writing, and speaking (as evident on the exam at first glance) copes with international recent research to date or not and whether this assessment method is still appropriate for adoption by investigating to date related advantages and disadvantages.

Assessment reflects what is taught. Teachers cannot teach something and perform partial assessment of what they have taught. Islam et al (2021: 1) states that assessment effectiveness is related to the method used in testing students' performance of the tasks being assessed and that any curriculum requires adopting the choice of effective testing methods.

Academically, it is proven that any assessment method should include various types of questions to handle various language components/skills. It is well known that not only multiple-choice but also other types of questions and practices can be well employed in assessing language, e.g. open-ended questions (Sychev et al, 2020), error correction (Khansir, 2018), true or false (Frost, 2005), portfolio-assessment (Arif et al., 2021), essay/ text writing questions (Sychev et al., 2020), as well as translation-performance/task-based translation questions (Alenezi, 2020), language situation ones/dialogues assessing language functions (Jennach et al., 2021), to maintain and cover all which could be included in an English exam type wise).

Limitations for the multiple-choice method is evident in literature. Although multiple-choice questions in particular can cover a wide range of topics to be assessed (Harris, 1969), they, on the contrary, cannot assess the productive skills of writing and speaking (Ibid: 7), as they do not require performance. It is also found in qualitative research that

senior teachers of English in Indonesia revealed that utilising multiple-choice questions in summative/final assessment had three advantages and three disadvantages (Rachmat, 2019: 10). The advantages include easy marking or scoring, assessment of students of varied abilities, and careful students answers. The drawbacks include “*low* order of critical thinking,” limited positive backwash and much time consumption in designing questions.

The most negative demerit of multiple-choice questions is related to cheating. Harper et al. (2020: 263) finds that this type of questions is easier to cheat by students, which is more difficult for invigilators to discover. The study also finds that essay questions are not as easy to cheat by students, but easier for invigilators to discover cheating on.

It is also seen that although multiple-choice questions do not need much time to prepare and score (Javid, 2014; Rachmat, 2019), they could fail to assess students’ skills if they are vaguely worded (Jannah et al., 2021; Jannah et al., 2021). It is also seen that while multiple-questions are rather objective and have a definite answer model (Weimer, 2018), they can be a failure if the questions are not well prepared (Jannah et al., 2021: 9). (This answer first set first question; see **1.6. Questions of the study and 3. Results and discussion.**)

The four skills of the English, i.e. listening, speaking, reading, and writing, from another perspective, need to be included in any assessment method in order to determine whether a student has met the educational goals in the language and whether teaching is done appropriately. Latha and Ravichand (2019:1) asserts that assessment of the four of the language skills feeds the subsequent teaching of those skills. Akongoh (2021:1) states that the speaking skill, in particular, needs to be assessed because of the role it plays in English language teaching, and because it provides “information on the effectiveness of teaching methodology, learner comprehension, students competencies and could be very useful for taking administrative decisions.”

Ibrahim and Othman (2021) carried out a study using *error analysis* to evaluate students’ writing and saw that students had to have the ability to write communicative texts and be assessed on the aim and content of writing. It is worth mentioning that the two productive skills (i.e. speaking and writing) are extremely needed for language production and are indispensable to be covered in any assessment method.

Listening, too, plays a key role in oral communication and contributes to developing other communication skills (Vu et al., 2022: 41) and Vu et al. carried-out research proving that a software called *Listening Hacked*, was used effectively in EFL listening assessment, as they believed in the importance of assessment to such a skill. Of course, listening needs *equipment* for playing it, if the teacher is not a native speaker. Putri and Susanti (2022: 14) indicate that reading comprehension assessment implies standardized real assessment, proving it has to be assessed at the end an educational bulk of time and used *Edomodo* as a tool for assessing reading comprehension. Thus, it is perceived that the assessment of such skills of listening and reading is required and needs to be effective.



Not only the assessment of language skills but also that of other linguistic elements (later called language elements) is needed. Recent research, e.g. Daneshfar et al. (2018), Kamali et al. (2018), Restaji et al. (2021), and Syvac (2018), deals with grammar assessment. Nugroho and Rekha (2020), (Jennach et al., 2021), and Echcharfy (2019) address language function assessment. Popovic (2020a) and Miralpeix and Muñoz (2018) deal with translation assessment. The following record screens those studies in more detail, to be briefed on up-to-date methods of assessment for the language components.

Definitely, there are recent, as well as old-rooted approaches, to the assessment of grammar as a language element. Daneshfar et al. (2018:295-297) find that *Dynamic Assessment (DY)* of grammar, i.e. which depends on assessing students' abilities which they can actually do unassisted and assesses those which are already in progress through interaction with the teachers or with other knowledgeable students.

It is evidently seen that DA clarifies the growing abilities of a students, and therefore, can tell of unseen/unassessed aspects of students' abilities. Daneshfar study (Ibid) found that learning with DA produced better results than learning with traditional/static assessment. These results are supported by Restaji et al.'s (2021: 1). The study reveals that there are near-end results DA with grammar achievement of EFL learners at intermediate level (Ibid). Kamali et al. (2018: 72) supports the view that DA can lead to student better results. Syvac (2018) also sees that all grammar elements in a course need to be covered on assessment.

As for language functions, they must be taught as they help in *real- world* communication (Nugroho and Rekha, 2020) and conversation exercises can assess them (Jennach et al., 2021). Echcharfy (2019: 361) advocates language functions forming communicative competence and leading to what is called interculturality learning showing people's understanding of each other's culture in a multi-cultural society. Therefore, it is contended they must be there at the time of summative assessment

Popovic (2020a) sees that *manual* translation, opposed to ready-made multiple- choice translation in the exam method, as will be discussed, can inform of error areas and can provide better educational feedback. Also, it is proven that vocabulary magnitude or number of pieces is seen as crucial element for students learning of language (Miralpeix and Muñoz, 2018) and thus related assessment.

Novel, too, is taught and assessed by the new method for the GCSE exam of English. It is agreed worldwide that the method assessing a novel component must handle a number of *elements* to be assessed. It is agreed worldwide that a novel as a genre of fiction has a number of elements. Burgess (2022) denotes that the novel components are character, plot, setting, narrative method or point of view, scope, symbolism, interpretation in life, entertainment. In common terms, elements of a novel represent character, setting, plot, narrative method, climax, and end. It has to be clear that any exam method needs to address students' understanding with respect to the previous elements as they form up the main ones.

Of this, it can be seen that all language skills and language elements need to be practiced and assessed in practice and covered by any assessment method. Students need to do four-skill performance-based tasks and be tested on language functions, translation, vocabulary and idioms. Also, it is seen that multiple-choice questions are a mixed-blessing, having merits and demerits, but the best merit is probably covering a wide range of topics, and the worst demerits are definitely the probability of cheating existence and hence, not reflecting students' real levels and the inability to assess actual performance of speaking, writing, and translation. It has also been found that multiple-choice questions are *low* on assessing higher-order thinking skills. It is furthermore been alluded that other types of questions can be used in the assessment of different language components, such as essay writing, error correction, portfolio, quizzes, besides multiple-choice. Moreover, it is concluded that novel has various common components, such as theme, characters, plot, setting, climax, and end, which need to be covered in any assessment method.

### **Context of the problem**

It is noteworthy that the M.o.E. adopted a new method of exam assessment depending solely on multiple-choice questions and using electronic marking of exam questions. This method abandoned what had been known for years as using various question types, such as paragraph questions, translation performance questions, and gap-fill questions. The M.o.E. method was also seeking GCSE electronic exams. It was furthermore noted the curriculum changed over the school years 2020-2021 and 2021-2022, but the method remained the same although there were voices asking for change.

The new M.o.E. new GCSE exam method, then, caused much argument in the Egyptian society in recent years, especially on Social Network Websites, i.e. Facebook, Twitter, WhatsApp, etc. Reasons for this included people's uncertainty if school infra structure was ready, including internet connection issues. Another reason was that students were not trained to answer questions the way the new method required, i.e. electronically or in the form of multiple questions solely.

Two of the researchers' children, a son and a daughter, passed through the GCSE exam and exam method. Fear was dominating at home. It was most prevailing when instructions were received from the ministry on a regular basis, as to the nature of the exams. The family was not certain about whether the exam would be electronic or traditional. The most fears arose from the exam being electronic, as the family new any mistake could cause the son or the daughter over a decade's waste of efforts. Also, other fears emanated from the nature of questions where there were only choices, with some of question alternatives vague/ill-worded with no right answer clear. There appeared to be no uncertainty about accurate model answers, especially when the ministry refused to declare an official model answer, which made people more uncertain and angrier about the new exam method.

Students' marks in English and other subjects were lower than ever before. Most parents appeared even dubious of the new exam method. They wanted their siblings to be well-paid graduates of the faculties of medicine, pharmacy, dentistry, engineering. All these faculties required high grades on the GCSE exams. All these observations,

implications, and thoughts led the researcher to wonder the efficacy of such a new exam method whether by means of an academic analysis of the method or in the perceptions of field stakeholders, such as teachers and students.

Besides, it was seen on the exam that the exam method did not consider listening and speaking questions, nor writing and translation tasks. These are considered essential components in a language course and must be assessed if performance and proficiency are considered. The lack of certain curriculum components was observed when looking at the exam, with the researcher being an academic associate professor of ELT. It was believed quite a thorough evaluation of the exam and exam method was necessary.

### **Problem of the study**

As seen earlier, it is noted that other types of questions and practices can be well employed in assessing the English language (Khansir, 2018), (Frost, 2005), (Arif et al., 2021), (Sychev et al., 2020), (Alenezi, 2020), (Jennach, 2021).

Based on the academic evidence reviewed, school unreadiness for change (Hasan et al., 2021), social setting argument and background, and the researcher's own experience and observations, there appeared to be field unrest as the GCSE exam and exam method appeared not to cater for the stakeholders' needs as well as research/academic evidence.

There was also an optimum-level research gap represented in the inexistence of national evaluation of GCSE English exam and of the newly-introduced related assessment method.

### **Purpose of this study**

The purpose of the present study was to 1.) evaluate the GCSE English exams and exam method of the school years 2020-2021 and 2021-2022 against the language skills and elements (language input) embedded in the English curriculum/ curriculum goals and 2.) investigate stakeholders' perceptions of the exam and exam method (i.e. those of teachers and those of students) and the perceptions of academic experts.

### **Questions of the study**

In order to attain the purpose of the present study, it sought to answer the following two sets of **questions**:

#### **First set:**

1. What are up-to-date academic advantages and disadvantages of the multiple-choice questions as forming the sole question type in the exam method?

#### **Second set**

1. How far do the 2020-2021 and 2021-2022 GCSE English exams and related exam method represent the language skills and language elements in the M.O.E. Curriculum?
2. How far does the exam and exam method for the school years 2020-2021 and 2021-2022 represent the novel elements in the two-year curricula?
3. What are the strengths and weaknesses of the new Exam method?



4. What are the 2020-2021 and 2021-2022 students' perceptions about their exams and exam method?
- 4.1. What are the 2020-2021 students' perceptions about their exam and exam method?
- 4.2. What are the 2021-2022 students' perceptions about their exam and exam method?
5. What are particular plans for improvement, if needed?

#### **Delimitation of the study**

This study was delimited to a number of teachers from the governorates of Ismailia and Sharkaya (representing six educational directorates from Ismailia and six from Sharkaya - see Table (1)) and a number of English Language Teaching (ELT) professors from both Suez Canal University and Zagazig University. The study was also delimited to a number of 1<sup>st</sup> year students enrolled at the Faculty of Medicine (F.o.M.), Suez Canal University, who attended the 2020-2021 GCSE exam and a number of students (F.o.E., 2<sup>nd</sup> year students) who attended the 2021-2022 English exam, both as GCSE graduates. It has to be clear that the study covered the exams of the school years 2020-2021 (Scientific Section) and the 2021-2022. It has to be noted that the exams for the school years 2020-2021 and 2021-2022 originated from the same question bank source set by the General Authority for Exams and Educational Evaluation affiliated to the M.o.E. and they followed the same question withdrawal rules. They both had the same number of questions related to a language input category. The procedures of this study took place two years after the new exam and method had been applied (July-October, 2022).

The study used the goals either stated or embedded in two different curricula for the two school years. The M.o.E. changed the GCSE curriculum from Haines (2018 to Chappell and Hart (2021a) and Chappell and Hart (2021b). A comparison was made between the goals and language inputs of the two curricula and the two exams to assess the degree of reflection of curricula in the exams.

#### **Significance of the study**

This study may be claimed to be of direct importance to the educational system in Egypt. It raises awareness towards necessary steps for revising GCSE assessment method in terms of GCSE Exam coverage of curriculum content and skills, question types used for assessment, adoption of performance-based forms of assessment, consideration of continuous assessment as an effective assessment means, and interest in stakeholders', experts', and students' views. Though the study is of relatively small scale, its implications may be considered quite representative of those in the original populations of teachers, experts, and students. Therefore, the study results may be generalized as participants came from population characteristics and circumstances representing the those of the source populations. The study may also, to an extent, be considered an attempt which fills research gap regarding the inexistence of a study evaluating the English national GCSE exams and the related assessment method.

As regards research, this study contributes to practical knowledge level-raising. It is vastly involved in providing invaluable knowledge to research in the filed of evaluation

of language assessment. Literature covered satisfies the need for several curriculum component coverages in related to appropriate and inappropriate assessment methods used and the recommendation section offers guidance on methods of language assessment to be used (see the **Conclusions and recommendations** section). Particularly, the section for literature reviewed in regard to advantages and disadvantages of multiple-choice questions (in connection with answering first set research question stated later) provides valuable advice for researchers and practice teachers.

## METHODOLOGY

The study adopted a qualitative approach to gather data and analyse results. It incorporated investigating teachers', experts', and students' opinions of the English exam and of the new assessment method in the two school years of 2020-2021 and 2021-2022. Reasons for this include facts that teachers are practitioners, observers, assessors, and, above all, stakeholders of the situation. Likewise, students are stakeholders who are affected by the exam. Academic experts, too, can inform of the situation and draw a course of action. The study designed and administered a questionnaire to be administered to both teachers and academic experts, another (in two versions) with students and an interview instrument with the teachers and experts at different phases of the evaluation.

### Participants

#### Teacher and expert participants

The study had *fifty* participant teachers of English from both Ismailia and Sharkaya Governorates (twenty-three from Ismailia and twenty-seven from Sharkaya), as well as *five* professors of education, three of whom are specialized in English Language Teaching (ELT) and two in assessment and educational measurement working at both Suez Canal University Faculty of Education (F.o.E.) and Zagazig F.o.E. Both the teachers and academic experts were jurors. The teachers were selected from Ismailia and Sharkaya as both governorates were much representative of the nature of the majority of other governorates and teacher populations in Egypt. Both share the characteristic that their communities are composed of government/private sector employees, farmers, and Bedouins, have the internet facility, and represent both modern and less privileged areas, exactly like most Egyptian governorates. Table (1) shows the teachers from both governorates worked at the following educational Directorates:

**Table (1) Educational directorates participating in the study**

<b>Ismailia</b>	<b>Sharkaya</b>
North Ismailia	Al-Heseinaya
South Ismailia	Fakous
Fayed	East Zagazig
Al-Tal Al-Kebeer	Abu-Hammad
West Quantara	Abu-Kebeer
Al-Qassassin	Hehia

The study also designed and used a questionnaire with GCSE graduate students of the school year 2020-2021 and 2021-2022. The study approach, then, combined three instruments and relied on qualitative analysis of data from teachers, experts and students.

The questionnaire for teachers and experts was sent to them via WhatsApp (the Social Website Application for social networking communication) and data were collected in the same way in July-October, 2022. Participants were required to fill two copies of the questionnaire asking for their view on the school years 2020-2021 and 2021-2022 exams, being sent the relevant copies of exams, and, in case of experts, student books as well. Participant teachers and experts took part in the questionnaire in order to answer the first study question mentioned earlier and they attended semi-structured E-interviews via the Zoom Application (i.e. to give insights into the first study question, and to provide answers related to the rest of study questions, except that for students' perceptions.)

### **Instruments**

#### **Teacher and Academic Expert Questionnaire (T.E.Q.)**

A questionnaire (Appendix 1) was designed and used to examine teachers' and academic experts' views regarding whether there was a match between the skills in the curricula of the 2020-2021 and 2021-2022 GCSE students and the actual representation of those skills in the final assessment via the new assessment method. Teachers and experts were requested to add their views regarding the new assessment method depending solely on one-type of questions for various skills and content topics. Again, the aim of this procedure was to help answer the first study question stated earlier.

#### **T.E.Q. design**

In order to design the questionnaire, a detailed analysis of M.o.E. curriculum language skills and element details stated at the beginning of M.o.E. English textbooks for both years and included in curriculum goals was first obtained. The analysis procedure of language skills and elements was necessary to identify which skills and elements which were needed to be assessed via the method in the exams of the two years.

For the language skills and elements in the M.o.E. English curriculum of 2020-2021 (Haines, 2018), the language input was reached as it was stated at the beginning of the Student's book (For an image of the language input goals/ language skills and elements, see Appendix 2). Appendix (1a) reflects the T.E.Q. embedding those skills and language elements in the 2020-2021 curriculum. The following Table (1a) reflects the total skills and language elements per area:

**Table (1a)****Numbers of language skills and elements the 2020-2021 Student's Book**

Skill/element	No.
reading	9
writing	13
listening	5
speaking	12
grammar	19
Language	15
functions	
vocabulary	85
idioms	17

For the curriculum of 2021-2022 (Chappell and Hart, 2021), which is seen similar to some extent, the skills and language elements were both either extracted or inferred from the Student's Book and Workbook (Chappell and Hart, 2021a) and Student's Book and Workbook (Chappell and Hart, 2021b) whether from the curriculum goal section or within the language input lessons themselves. Sometimes when there was no separation of skills in the curriculum goal section of the Student's Book and Workbook (only one book for both content and practice of this content); then, the skills were extracted and included in the questionnaire to examine participants' opinions. This procedure was done by the researcher in order to identify the language skills and elements needed for evaluation to see whether they were covered by the exam or not (For a detailed list of language input/ skills and elements, see Appendix 3).

Again, it is noteworthy that not all skills were not explicitly stated in the Student's Book and Workbook (Chappell and Hart, 2021a) and Student's Book and Workbook (Chappell and Hart, 2021b) also because the Teacher's Guide mentioned that it adopted an integrated skill approach where more than one skill was presented together with a topic (Longman, 2021: 2) and this was reflected in the Student's Book and Workbook.

A whole analytical survey of the language skills and elements was made from the Student's Book and Workbook (either mentioned in the goal section of the Student's Book and Workbook or derived from the lessons by the researcher). Table (1) shows the numbers of those skills and elements.

**Table (1b)****Numbers of language skills and elements the 2021-2022 Student's Book and Workbook**

Skill/element	No.
reading	12
writing	12
listening	3
speaking	15
grammar	17
Language	1
functions	
vocabulary	157
idioms	23

Table (1b) above displays the total numbers of skills and elements in the curriculum. As indicated in the table, there are twelve reading, twelve writing, three listening, and fifteen speaking skills. It also shows that the language elements are seventeen grammar (i.e. topics), one language function, 157 vocabulary items, and twenty-three idioms.

Through the analysis of M.o.E. curriculum Chart of Skills (under the goals section) at the beginning of the Students' Book and Workbook, it was clear that the language input was merely in the shape of language input mingled with the topics: grammar topics, including one language function, and no vocabulary-related elements, and no idiomatic expressions (i.e. quite contrary to what research advises regarding informing students what they are going to learn beforehand, e.g. Latha & Ravichand, 2019). The lesson surveys revealed that vocabulary and idiomatic expression teaching were found in the lessons themselves, being not evident in the Chart of Skills. It was a *remark* that the vocabulary elements and idiomatic expressions were not explicitly listed in the Students' Book and Workbook. The lesson surveys included a manual calculation procedure of vocabulary and idiomatic expressions in the lessons, being unidentified in the Chart of Skills, in order to determine the number of words and idioms the lessons contained. 157 vocabulary items and twenty-three idiomatic expressions were found across the whole lessons, as previously indicated in Table (1).

The design of the questionnaire (two versions suiting the two-year curricula as in Appendix 1a and Appendix B) had all the skills and language elements in the Students' Book (Haines, 2018) for the school year 2020-2021 and Students' Book and Workbook (Chappell and Hart, 2021) for the school year 2021-2022. They were shown as and divided into language skills section: reading, writing, listening/speaking skills; and language element section: grammar, language functions, vocabulary, idiomatic expressions, translation, and novel-related elements. The questionnaire was made in two versions to cope with the change in curriculum over the two year. The two-version questionnaire had both 'fixed alternative' to see if the skills and elements were represented/covered by the exam or not. The participants were required to state whether a skill or element was represented on the exam by the method through two-scale fixed alternative choices of either 'Agree' or 'Disagree.'

### **Statistical considerations**

Special attention, i.e. in regard to the nature of data calculated by the questionnaire in its two versions and treated statistically, needed to be paid. It was necessary and justifiable that in final calculation of curriculum skill counts for the questionnaire, all repetitive skills in the curriculum/curriculum goals were dealt with as *one* skill and the same procedure was performed across and within all skill areas (the biggest macro-skills), and skills joined together, such as skimming and scanning in the reading area, and making introduction, body, and conclusion in the writing one, were broken down and calculated as the number of sub-skills they contained (two and three, respectively in this case).

### **T.E.Q. Validity**

To check the validity of the two-version questionnaire, as in Appendix (1a) and Appendix (1b) for the two consecutive years, Face Validity was conducted. It was sent



via the internet to six full-time professors of ELT from both Suez Canal and Zagazig Universities. Four of whom were specialized in ELT; two in Assessment and Educational Measurement. Responses were collected and related modifications were made regarding necessary questionnaire separation of skills for the version related to the exam year 2021-2022 as they were sometimes integrated with content as in the Student's Book and Workbook (Chappell and Hart, 2021a) and Student's Book and Workbook (Chappell and Hart, 2021b).

### **T.E.Q. administration**

The questionnaire instrument had both the skills and language elements which were obtained from the books of the two consecutive years (Haines, 2018), which has a separate workbook, and the Student's Book and Workbook (Chappell and Hart, 2021a) and Student's Book and Workbook (Chappell and Hart, 2021b). Both the informal copy of the formal exams and M.o.E. curricula for the two years in PDF formats were sent to both the teachers and the experts. They were asked to see both the skills and the exam questions in order to decide whether they agree/disagree that each skill/ language element was assessed by/covered in the exam via the exam method. The two-version questionnaire, i.e. as in Appendix (1a) and Appendix (1b) for the two consecutive year exam was a two-point verdict scale (Agree/Disagree) for participant decision whether a language skill/ language element was represented on the exams, as this was what was exactly adequate.

### **Examination of the GCSE Multiple-choice Exams**

The exams used for examination and analysis by the teachers and experts were those of the school years 2020-2021 and 2021-2022. They were informal copies of the formal exams as the M.o.E. has not uncovered formal copies yet. (It was noted that it both exams may be seen comparable at first sight as they still shared the main format and multiple-choice essence, and total number of questions and, as listening, speaking, and actual writing and translation were common missing features at first glance in both exams and they utilised the same method). For the 2020-2021 exam, see Appendix 5 and for the 2021-2022 one, see Appendix 6.

The exam total mark for the two school years was fifty, with total questions of forty, which meant ten questions received two points. The M.o.E. mentioned that highest-thinking questions deserved double points. All questions were multiple-choice. At first sight, the exams had no questions for listening and speaking. No performance tasks of any type were required on the exam.

During the interview with the academic experts of assessment and educational measurement, it was noted that the experts occasionally tended to avoid answering a number of particular questions related to the specifics of the English language although they answered all other questions. They thought ELT professors were in a better position to give direct views on the language as their major was not English-related, although it is education-related. Each time they withdrew was recorded, and the other experts were only included in view calculation.

### **Interview**

To be able to confirm, analyse, and discuss the answer of the second set first and second study questions, being primarily obtained by means of the fixed alternative questionnaire, and to answer both second set the third and fifth, the interview with both the teachers and the experts was performed in two phases: one with teachers and another with experts. The interview results were used to have more meticulous insights into and occasional support to the answers of the T.E.Q., to obtain a confirmation of the answer to the second set first and second study questions mentioned earlier, and to obtain clear answers and interpretations of the second set other study questions in relation. The interview consisted of seven questions:

1. Does the assessment represent/match the targeted the skills included in the textbook?
2. Does the exam cover all the skills and language elements presented to the students during the school year? How far is the skill/language element coverage? What skills/elements are missing?
3. Do you think there is a need for the inclusion of other subject-related types of questions, such as, actual essay, email, text message, etc. writing, actual text-translation, and more narrative questions related to theme, character, plot, etc.?
4. Do you think that new assessment method can help with the assessment of all language input? How helpful is it – not helpful, moderately helpful, or much helpful?
5. Do you think the Novel, *Great Expectations*, the students have taught is well represented in the final assessment? Why/ Why not? If any, how far is the representation?
6. What do think of the 2020-2021 and 2021-2022 English exams and related assessment method? Identify aspects for strengths and weaknesses, please.
7. What do think you need to add for improvement as a practitioner teacher/academic expert?

Questions number one and two in the interview dealt with confirming and supporting results of the second set first study question. Question number 5 was used to answer the second set second study question. Question numbers 3, 4, and 6 were used to answer the second set study third study question. Question number 7 was used to answer the second set fifth study question.

### **Student questionnaire (S.Q.)**

To obtain student perceptions of the GCSE exams of the school years 2020-2021 and 2021-2022 in order to answer the fourth study question (4.1. and 4.2), a questionnaire was designed, validated, and administered to graduate GCSE students who attended the English exams of the school years 2020-2021 and 2021-2022. The S.Q. included ten questions assessing various exam and assessment method aspects, such as coverage of all what the students had studied, the topics which were not covered, likes and dislikes about the exam, challenges on the exam, method of exam adoptions of various question types (For the detailed questions, see Appendix 4).

### **S.Q. validity**

An ELT professor from F.o.E., Suez Canal University, together with the researcher, designed and Face-validated the questionnaire. A means of telephone discussion on appropriate meaningful questions suitable for students' language levels and understanding was followed in order to reach agreement on the final questions to be used in the questionnaire. The Face validity was established.

**Administration of the S.Q.**

The S.Q. was sent via the Social Network WhatsApp Application to 100 GCSE graduate students who attended the 2020-2021 exam and 100 GCSE students who attended the 2021-2022 exam. The former students were selected from Faculty of Education at Suez Canal University, and those who attended the 2021-2022 exam from the F.o.M. Selecting certain faculty participants was not purposeful, but the selection reflected what as possible for the research. Results was obtained the same E-method of data collection. Responses were collected in October 2022.

**RESULTS AND DISCUSSION**

Data collected by the three instruments much aided, not only in question answer confirmation purposes, but also in identifying the extent to which the final M.O.E. exams represented the skills and elements implied in the M.O.E. curriculum and obtaining the perceptions of teachers, experts, and students. Results of the three data instruments were screened, analysed, and discussed. First, data from both the T.E.Q. (in its two versions for the two consecutive years) and the interview were statistically and qualitatively analysed. Then data from the S.Q. was addressed in the same way. The following section deals with the results (obtained by the T.E.Q. in both versions and related interview information) and addresses the study questions, related analysis, discussion, and interpretation.

**Answer to the first set questions:**

In answer to the first set question, “What are up-to-date academic advantages and disadvantages of the multiple-choice questions as forming the sole question type in the exam method?”, it is clear that this question was answered during the course of the literature reviewed. See 1.2. **Multiple-choice assessment method.**

**Answer to the second set questions**

In answer to the Second Set first study question, ‘How far do the 2020-2021 and 2021-2022 GCSE English exams and related exam method represent the language skills and language elements in the M.O.E. Curriculum? The T.E.Q. and first two interview questions addressed the answer in lengthy detail.

As for relating T.E.Q. data for the 2020-2021 English exam, **Table (2)** shows participant positive or negative responses and positive/ negative response percentage.

**Table (2) Teacher and academic expert responses and related percentages**

	No	Percentage
<b>Number of teachers stating, “Yes”.</b>	5	10%
<b>Number of teachers stating, “No”.</b>	45	90%
<b>Number of academic experts stating, “Yes”.</b>	0	0%
<b>Number of academic experts stating, “No”.</b>	5	100%

Table (2) indicates that 45 teachers did not agree that the exam represented all the skills and language elements in the curriculum. Also, it shows that five experts of five had

the same views. This is quite clear as they (n. 50 for teachers and 5 for experts) confirmed in the T.E.Q that listening and speaking missing, and writing and translation performance were not present, either. This was a straight forward first remark, as stated earlier. The T.EQ. also shows more specific details related to coverage of the skills by area and language elements by the exam.

**Table (3)** shows number, and related percentage of, exam questions (E.Qs.) covering a whole area, related percentage (P.) in relation to the whole exam questions, number of skills (Ss.)/language elements (L.Es.) covered in that area by those questions, percentage of Ss./L.Es. covered in relation to the area, number of Ss./L.Es. missing in assessing the area, percentage of Ss./L.Es. missing in relation to the area and related percentage to area Ss/L.Es., and frequency of questions covering the same skills/Ss./language element/L.Es. in the exam.

**Table (3)**

		No. & P. to E.Qs.)	No. of Ss./ L.Es. cvd.	P. to all Ss/L.E s. in area	Ss./L.E s. missin g	P. missin g Ss/Ls. to area	F. of qs. for the same Ss./L.E.
<b>Skill</b>	Reading	16 (0.40)	3	%33	6	%67	0
	(Knowledge about) writing	4 (0,1)	1	%7	12	%93	-
	listening	0(%0)	0	%0	5 (all)	%100	-
	speaking	0(%0)	0	%0	12 (all)	%100	-
<b>Total</b>		20	4	-	47	-	0
<b>Linguistic Element</b>	Language/ grammar	10 (0,25)	10	%48	10	%52	0
	Language/ vocabulary	6 (%15)	5	%4	150	%96	-
	Idiomatic expressions	0 (%0)	0	%0	17	%100	-
	Language Functions	0 (%0)	0	%0	1	%100	-
	Language/Translation	4 (%10)	4	%100	%0	%0	-

As indicated in Table (3), 16 (40%), 4 (1%), 0 (0%), and 0 (0%) questions for reading, writing, listening and speaking, respectively were on the 2020-2021 English exam out of a total of forty exam questions. Only three skills for reading, one for (knowledge about) writing, and no questions for both listening and speaking were covered. With reference to Table (1a) total skills in the curriculum, this forms the following percentages to all skills in the curriculum: 33%, 7%, 0%, and 0% for reading, writing, listening and speaking, respectively. Table (3) also shows that the missing skills in the four areas, considering the totals in Table (1a), are 6 (67%), 12 (93%), 5 (100%), and 12(100%), respectively. The table also displays that there were no redundant questions for the same skills on the exam.

Table (3) further shows that the exam questions representing grammar, vocabulary, idiomatic expressions, language function, and translation were 10 (0,25), 6 (%15), 0 (%0), 0 (%0), and 4 (%10), respectively, out of a total of 40 exam questions. With reference to Table (1a) displayed earlier, Table (3) also shows that the elements

represented on the exam were 10, 5, 0,0, and 4 for grammar, vocabulary, idiomatic expressions, language function, and translation, respectively. There were no repeated language element questions.

This informs of the incomplete representation of both language skills and elements of the exam. Only reading, grammar and translation can be seen representative to an extent, while writing, listening, speaking, idiomatic expressions, and language functions were completely missing or severely unrepresented on the exam while vocabulary was inadequately represented. This might be attributed to the nature of the method itself which does not support the performance of speaking and actual writing but rather assess knowledge about it through multiple-choice questions. There were no obvious reasons why other language elements were not at least mostly represented.

As for the T.S.Q. data for the 2021-2022 exam, both the teachers (n. 47/ almost all) and experts (n. 5/all) showed that not all the skills were represented to a considerable extent. The following Table (4) shows number and related percentage of exam questions (E.Qs.) covering a whole area and related percentage (P.) in relation to the whole exam questions, number of skills (Ss.)/language elements (L.Es.) covered in that area by those questions, percentage of Ss./L.Es. covered in relation to the area, number of Ss./L.Es. missing in assessing the area, percentage of Ss./L.Es. missing in relation to the area, and frequency of questions covering the same Ss./L.Es. in the exam.

**Table (4)**

		No.& P. to E.Qs.)	No. of Ss./ L.Es. cvd.	P. Ss. to area	Ss./L.Es. missing	P. missing Ss. to area	F. of qs. for the same Ss./L.E.
<b>Skill</b>	Reading	15 (0,37)	6	%50	6	%50	2
	(knowledge about) writing	3 (0,1)	3	%40	4	%60	-
	listening	0(%0)	0	%0	3 (all)	%100	-
	speaking	0(%0)	0	%0	13 (all)	%100	-
<b>Total</b>		18	9	-	26	-	2
<b>Language Element</b>	grammar	11 (0,26)	7	%53	10	%47	3
	Vocabulary	7 (%18)	7	%0,04	150	%100	-
	Idiomatic expressions	0 (%0)	0	%0	23	%100	-
	Language functions	0	0	%0	1	%100	-
	Translation	4 (%10)	4	%100	%0	%0	-

Note that all percentages are approximated to the nearest unit.

Table (4) shows the skills and language element areas which were represented by/ missing on the 2021-2022 exam. The table demonstrates that the numbers of questions used for the areas of reading, (knowledge about) writing, listening and speaking were fifteen, three, zero, and zero, respectively, out of forty total exam questions, with total twenty-six skills missing in the four areas out of thirty-three. Relating percentages of



the four-skill representation/coverage on the exam show *very* low levels for writing, listening, and speaking representation on the exam. As shown above, they were 40%, 0%, and 0%, and 0% respectively. The table also shows that reading received a representation of 15 questions worth of %50 skills representation, being the highest represented.

The data in table (4), moreover, indicate that a big number of skills are either inadequately-represented (%from 40 to %50) or completely missing (sheer %100 for both listening and speaking) in the four skill areas. Therefore, this may cast big doubts on the M.O.E. practice enhancing students to adopt communicative oral skills of speaking and listening, in particular, as they are the skills unrepresented on the exam. The representation levels of both reading and writing are inappropriate, with writing the third less represented - (%40)/ missing (%60).

It is noteworthy that the M.O.E. Teacher's Guide by Longman (2021) states that the curriculum objectives are to develop many skills included in the teaching of the curriculum. This shows the *contradiction* between the curriculum goals and actual practice of assessment and related *imbalance*. It is necessary that the goals of a curriculum and its assessment be representative of each other (Latha and Ravichand, 2019), as clarified earlier.

The low number and percentage, 3 (0,1) of questions representing the area of writing assessed on the exam does not support proper assessment measures and coverage of material learned, nor does it reflect what the M.O.E. announces that it wants to see students having adequate skills by including seven genres in the curriculum, i.e. integrated with skills (see Appendix 3). Ironically, the Teacher's Guide (Longman, 2021) and the Student's Book and Workbook (Chappell and Hart, 2021a) and Student's Book and Workbook (Chappell and Hart, 2021b) imply that the teaching of writing essays (three types: persuasive, argumentative, and opinion), reports, emails, C.Vs., surveys, webpages, and text messages are included in the curriculum. For all genres integrated with the writing skills supposedly taught to the students, see table (1). Six genres out of seven were reported to be missing on the assessment during the interview. The biggest majority of seven genres important for today's life (see Table (1)) are missing and this does not cope with the four-points of OBE Spady (1994) clarifies.

Generally, when it comes to the what the two exams of 2020-2021 and 2021-2022 assessed as for writing, i.e. one genre of essay writing, it is found out that related questions only assessed a general type, not a *specific* type, of essay writing. They do not assess the very types of persuasive, argumentative, and opinion essay writing as instructed in the Teacher's Guides, but, rather, a general type. This can be considered a waste of teachers' efforts and time. As reported in the interview by the experts and teachers (n. 5 and 45 for experts and teachers, respectively), the missing types included a persuasive essay, opinion essay, argumentative essay, a report, online comments, email, invitation, survey, in the 2021-2022 exam while the same types were missing, excluding and argumentative essay and an invitation which were not included in the related curriculum. The experts (n 3) added a comment on both exams that "The exam lacked writing performance and questions were related to knowledge about it, not actual

writing.” A juror expert stated that “the M.o.E. could have assessed writing and still marking questions electronically if they had included ‘scrambled sentence or paragraph questions and asked students to reorder them to form the right genre format.’

Besides, it is confirmed that the sole nature of multiple-choice questions may *not* help with actual application of the writing process which necessitates brainstorming, planning, drafting, writing, editing, and revising (Kitjaroonchai, 2022; Seow, 2002;). The exams evidently do *not* require students to show these important abilities/processes.

For assessment purposes, it appears to be necessary for students to show their ability to practise the act of writing. However, the exams did not ask students to do actual writing and this was *not* the practice that students were supposed to be doing to improve their writing. The one skill practised for the 2020-2021 exam, the three skills practised for the 2022-2022 exam, as reported by the ELT experts (n 3), were only related to essay writing development for the former and the latter, and format and punctuation for both, although punctuation was *not* expressed in the introduction to the 2020-2021 curriculum nor in the 2021-2022 curriculum goals, even on an implicit level. However, it is found in both curriculum language inputs. It is worth noting that four and three questions for the two consecutive exams (almost %1 of the whole exams) for writing are extremely *inadequate* for assessing this important skill. Nor is 7% and %40 exam representation of writing skills on both consecutive exams acceptable; for these percentages, see Table (3) and Table (4) located earlier. Writing should be practised and assessed thoroughly (Ibrahim and Othman, 2021).

With respect to the skill of reading in Table (4), it is found out that the coverage percentage in relation to the total number of skills was *average* (%50). It has to be higher than this in order to have better insights into and verify the achievability of what has been taught. Reading, too, needs to be well taught and assessed (Putri and Susanti, 2022: 14).

Comparing the reading skill representation percentage of the 2020-2021 exam to that of the 2021-2022 exam, it is seen that both are 33% and fifty, which informs incomplete representation. Reasons for this might refer to the little time allocated for the exam (only one-time three hours). To be able to assess all skills, time reconsideration appears a necessity.

As shown in Table (4), even grammar elements received *inadequate* representation in relation to the total language elements supposedly taught to the students. The table shows that %47 of the grammar elements (in the shape of topics) were missing on the exam. This means that only %53 of them were covered.

Comparing this result with the related one for the 2020-2021 exam in Table (3), it is seen that 48% of the grammar elements (topics) were represented. This means that the two percentages are comparable and incomplete representation is thus inferred.

Nonetheless, it is well known that in order for students to structure their speech and writing, they need all grammatical elements to be covered and proper level obtained (Syvac, 2018), as mentioned earlier. The inappropriate percentage of grammar representation in the exam is seen as definitely contrary to what is stated in the M.o.E.'s teacher's Guide itself (Longman, 2021: 4) that one of the goals behind teaching this course is to get students to acquire a solid knowledge of the linguistic systems of English."

Unfortunately, as shown in the Table (4), three grammar questions on the exam reflected the same language element/topic. It was noticed that one referred to advanced feature of 'articles', another for past perfect, and the third for relative pronouns. However, it is thought that redundant elements should have been replaced with unrepresented ones, such as *present perfect continuous*, *would for past habits*, and narrative tenses, for example, according to all participant teachers and experts during the interview. More questions representing grammar should have been introduced, as reported by the majority of participants (43 teachers and four experts), as appraised during the interview. Also, two reading skills were repetitive, as shown in Table (4). The participants (all teachers and experts) made clear that this entailed a considerable number of questions to be repetitive, too, and they should have been used to cover skills missing in another area. It is worth mentioning that redundancies did not exist in the 2020-2021 exam.

Table (4) above also demonstrates that the language functions on the exam were represented by questions on the exam although one element was found in the objectives of the curriculum (i.e. textbook). Comparable to this, data from Table (3) indicate that no language function was assessed while there was a total of fifteen language functions in the curriculum as in Table (1).

It has to be voiced out that this is *not* good practice (and has to be *changed*) because recent research evidence to date sees that the success of intercultural communication requires grammar ability and suitable use of language functions within a social context in reality (Nugroho and Rekha, 2020: 2). It is, therefore, seen that teachers should arrange situations for language use (including various language functions such as offering, requesting, suggesting, taking permission, etc.) with native speakers or at least with a competent teacher/ student in order for students to practice them. Recent research also states that learning language functions should be directed towards what Echcharfy (2019: 361) calls an intercultural approach to language learning where the learners are exposed to different examples of language functions boosting communicative competence.

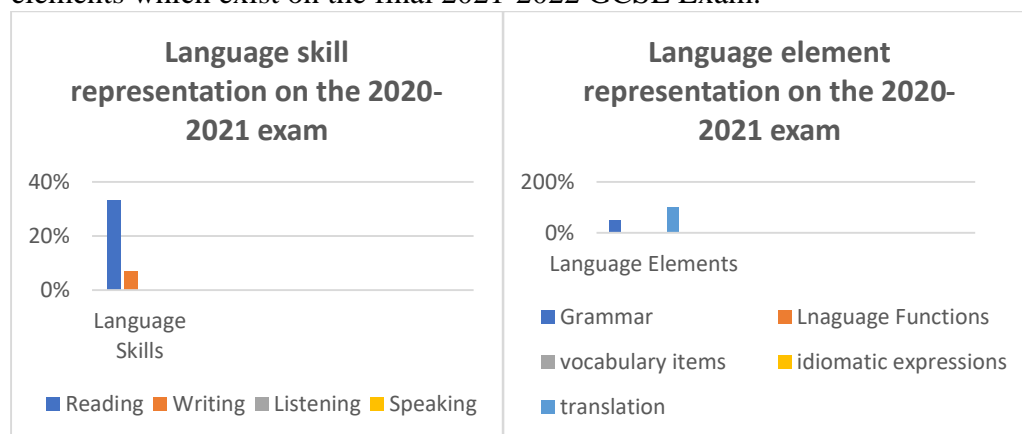
With respect to the vocabulary questions on the 2021-2022 exam, it has to be mentioned that, although they appear to be quite of representative in number, they are obviously of an unrepresentative nature. That is, although they make a percentage of %18 of the exam, as shown in Table (4), it has to be related that teachers (n. 39) on the interview said that most vocabulary they taught did *not* show on the exam. This is quantitatively analysed in Table (4) where 150 items appear to be missing on the exam (almost %100). Compared to these results are those related in Table (3). Six out of 85 items were only

assessed. This means that both exams lacked a wide and various range of vocabulary testing.

Idiomatic expressions were absent on the 2021-2022 exam as well (23 items/%100). See Table (4) above. The teachers (n. 41) interviewed were angry about this. See following interpretation about idiomatic expressions. The same happened on the 2020-2021 exam, a complete misrepresentation

As Table (4) shows, four translation questions (%10 of the whole exam question) were used. Th same happened for the 2020-2021 exam, as in Table (3). This is seen quite a good quantity representation. However, since multiple choice questions do not *allow* students to engage in *actual* act of translating texts. This cannot be judged as fair enough for the quality of questions as actual engagement in the act of translation by students is indispensable for improvement and as their errors in translation becomes clearer for instructors to help later educational decisions (Popovic, 2020a; Popovic, 2020b) and, thus, actual translation needs to be assessed for students' benefit and for teachers to obtain feedback.

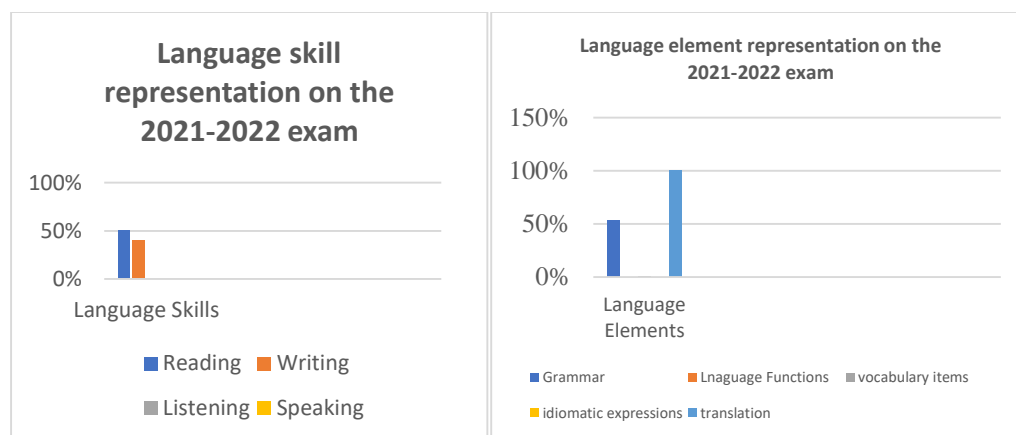
The following **Figure (1)** and **Figure (2)** represent the language skills and language elements which exist on the final 2021-2022 GCSE Exam.



**Figure (1)**

**Figure (2)**

The following **Figure (3)** and **Figure (4)** represent the language skills and language elements which exist on the final 2021-2022 GCSE Exam.

**Figure (3)****Figure (4)**

The more details from the interview results with respect to the second set first study question, confirmed the questionnaire results. All of the participants (i.e. 50 teachers and 5 experts) interpreted that the listening and speaking skills were *utterly* absent on the final assessments/exams and only grammar and vocabulary (language) were present to a *considerable* extent for both exams as they and the experts (n. 50 and n. 5, respectively) stated they shared common characteristics. This might be attributed to the assumption that the exams did not focus the attention on the productive skills of the language, nor did they cater much for listening which was observed on visits to schools and judged as rarely taught although its sources and activities were provided in the curriculum on a theoretical basis.

It is well known that speaking is also inevitable to language learning (e.g. Akongoh, 2021), and so is listening (e.g. Vu et al., 2022). Besides, it is well known that a language cannot be developed without those two communicative skills. Communication in everyday life in English would be quite impossible without listening and speaking.

The *practice* of writing on the exams was not assessed, either. Both the teachers and experts (n. 50 and 5, respectively) agreed that students on both exams were not required to write an essay. They agreed that there were a number of knowledge questions related to the understanding and application of one type of genre: essay writing. However, they stated in their lengthy interview comments that the following nine skills were *not* included in the assessment of both exams:

1. Report forming
2. Producing of an informal email offering help to a friend
3. Forming and using language appropriate for email
4. Making event invitation
5. Creating essay expressing opinion
6. Forming a survey
7. Structuring a text

Also, they saw that, besides those above, the 2021-2022 exam in particular did not address the following skills related to the reflection of online-comments and 2.the design of online career profile.



This meant that nine of the skills out of twelve ones were not assessed on the 2021-2022 exam, which casts great doubts on the validity of the exam (bearing in mind that both listening and speaking skills were not assessed at all.) Also, the participants (n 44) mentioned that even when two questions were related to essay writing, this did not mean students can write an essay well as the questions did *not* require writing, so “we knew nothing about students’ ability in writing.” “Most of the genres we taught like email, survey, webpage, etc. did not appear on the exam,” they added. It is worth clarifying that Latha and Ravichand (2019: 5) state that exams should reflect what is there on the “the stated outcomes of education.”

Compared to the skills missing on the 2021-2022 exam, those on the 2020-2021 exam were twelve out of thirteen, which is seen extremely unrepresentative. Possible causes can be referred to, again, insufficient exam time and independence on continuous assessment as a means, as discussed earlier.

Based on the above-based evidence, inferences from the two exam evaluations can be made. As common features, both showed the negligence of assessing the practice of three important skills, at least in the present study, a hazard which needs to be treated on an immediate basis (for possible interventions, see the ‘Results and recommendations’ section). On both exams, actually, the students were *not* asked to listen, nor to speak, nor to write, and nor to do the act of translation. Vocabulary was much underrepresented and idiom and language functions were completely missing. Language functions were not clear in the questions/ missing. Reading disregarded most of the reading types in the curricula though related questions on the exams formed considerable numbers. Thus, it may be commented that this may not be an appropriated practice to represent an exam of a country.

The three major/macro language skills mentioned above have the majority of language skills. The Teacher’s Guide for the 2021-2022 curriculum (Longman, 2021:1) declares that “it [the curriculum] aims to assist students in the process of teaching certain behavioural and civic proficiency goals not only in the English language but also in the day-to-day interactions which they will encounter though their lives.” The optimum query here is how would the ministry make sure that students’ proficiency and day-to-day interactions have been achieved without proper assessment? Some interpretation to the clearly negative answer lies in the presumption that the M.o.E. has not practically taken *measures* towards the adequate *teaching*; thus, assessment of listening and speaking, and even writing is thought not to have ample time for practice in the classroom. Thus, this negative practice unfortunately may reflect great assessment drawbacks as far as the final exam is concerned.

Also, during the interview, both the teachers and experts (n. 50 teachers and n. 3 ELT experts) confirmed that the 2021-2022 exam also had *vocabulary* questions – seven whose elements/items were absent on the M.o.E. so called *chart of skills* at the beginnings of the Students’ Book and Workbook of the two terms. (Actually, 157 vocabulary items and twenty-three idiomatic expressions were found and calculated, i.e. in the present study, from the lessons, as stated earlier, as in Table 1b). This might suggest the possibility of the random assessment nature of important language

elements, being neglected/ unclearly demonstrated at the introduction phase of the curriculum where goals are stated. The teacher's Guide for the 2021-2022 exam (Longman, 2021: 7) only gives general perspectives on informing the teachers of the vocabulary to be taught, stating: "New vocabulary items are introduced in the context of reading texts where vocabulary items are highlighted in clear bold font...The vocabulary is practiced in different activities in both sections of the Student's Book and Workbook." Obviously, the vocabulary items, as well as idiomatic expressions, need to be stated clearly in the goals of the curriculum.

In comparison, the 2020-2021 Teacher's Guide (Haines, 2018) had quite similar language input skills and elements, except that it was seen that the language functions were much more than the 2021-2022 curriculum. The 2020-2021 exam had six questions assessing vocabulary (Table 3) out of 85 vocabulary items (Table 1b). six and seven vocabulary questions for the 2020-2021 and 2021-2022 exams, respectively are much underrepresentation of the overall items, 85 and 157 for both consecutive year exams. Once again, three-hour exam in one sitting is seen inadequate for assessment of whole year student work. Continuous assessment has to be intervened, as suggested earlier.

The negligence of proper vocabulary and idiomatic expressions assessment on both exams, as displayed in Table (3) and Table (4), might genuinely be attributed to the assumption that the M.o.E. had not summoned F.o.E. experts for developing not only vocabulary-related but also other area assessment necessary for GCSE students. It is noteworthy that vocabulary magnitude is important to students (Miralpeix and Muñoz, 2018) and thus related assessment.

Also, it is commented that M.o.E. abandoning important language elements in the introduction of curricula might lead to unclear vision on the part of the teachers teaching them. Students could have been *victims* of such a negative practice, as teachers (n.50) said in the interview that "we were tired of focusing in our presentations/ explanations on idiomatic expressions to the students." (They mean it in private tutoring, not at schools where students conversely do not go) "None appeared on the exam," they added "Ha!Ha!Ha!" they laughed. (This might give grounds for *parents'* unassessed dissatisfaction disseminated in the social media about the final exam and caused by the sole multiple-choice exam method. Their anger appears to be justifiable.)

During the interview, the teachers (n. 50) also revealed that some of the reading skills, the skimming and scanning ones of two reading articles on both exams, were focused on one type of readings of reading out of seven taught. "They were all only one essay type," they explained the nature of the two texts which appeared on the exam. (For a complete list of the types of reading in the 2020-2021 curriculum, combined with the skills, see Appendix (2), and for those in the 2020-2021 curriculum, see Appendix (3). It is also worth noting that the reading questions on the two exams lacked timing as they were given with other language skills and language element questions, according the comments of one expert. Timing is needed for the skills of skimming and scanning (Fauzi, 2022: 101). The participants (50 teachers and the three ELT experts) said the exam contained skimming and scanning questions with much focused on those skills,

which may be seen fair enough by the present study. However, they said that fifteen out of forty total exam questions (%0.38), as shown in Table (4), assessed reading and sixteen out of forty (%40), as shown in Table (3). This is seen as a little too many questions, as there should have been and should be a *balance* in assessing the other three skills language macro-skills, bearing mind assigning weights for other language elements as well. The experts (n. 5) had the same views.

To determine what reading skills were missing on the 2020-2021 exam, both the teachers (n. 48) and the experts (n.5) stated during the interview that the following reading skills related to the following reading abilities did not receive *any* weight in the final assessment:

1. Summary-forming
2. Poetic figures-of-speech extraction
3. Poetic style judgement
4. Student CV-related language and section inference
5. Text type identification

To identify what reading skills which were missing on the 2021-2022 exam, both the teachers (n. 47) and the experts (n.5) stated on the interview that the following reading skills did not receive *any* weight in the final assessment:

1. Comparing dissimilar styles of newspapers
2. Text-summarising
3. Skimming and scanning text in science
4. Identifying language expressed in text messages
5. Recognizing author purpose for and form of a report
6. Identifying webpage main idea and details

This may exemplify a deficiency, as well as imbalance, in the final assessment of students' reading, too, as it should reflect all the skills taught. Six types of reading/texts are unfortunately missing ones in the 2021-2022 exam: a *newspaper article*, a *scientific text*, a *text message*, a *report*, a *webpage*, and a *C.V*. These types of reading appear to necessary in everyday and business life.

The teachers (n. 48) and experts (n. 3 ELT experts) clarified on the interview that the types which are missing in the 2020-2021 exam are four. These include poem, C.V., part of a book, and description text. Even, two ELT jurors stated that both of the exams assessed only literal reading by requiring only skimming and scanning of information and there was deep understanding or what is beyond the lines of texts, or author's purpose of writing.

The teachers (n. 47) in the interview argued that the text translation skills were absent in both the 2020-2021 curriculum introduction and the 2022-2021 curriculum *skills/language chart* provided by the M.o.E. and on the final assessments, except for four multiple-choice questions "which could not *thoroughly truly* tell of a student's level in translation on the two occasions." It is worth mentioning here that all the participants (teachers n. 50 and experts n. 5) expressed their *dissatisfaction* about the way translation questions were on the exams; obviously, they wanted the students to *do* the action of translating texts and write them and *show* their abilities in this area,

whether from English into Arabic or vice versa rather than choose or guess the right answer from four alternatives. The experts (n. 5) supported the teachers in their reflections in this respect.

During the interview, the participants (47 teachers and 3 ELT experts) saw that certain grammar elements/topics were not represented on the 2020-2021 exam. These include those related to the following grammatical structures:

1. Past perfect continuous
2. Future perfect
3. *Will / going to* for the purpose of prediction
4. Relative clauses
5. So, or such ... that, enough or too ... to
6. Causative verbs
7. Reported statements
8. Passive forms
9. Conditional forms
10. Sentences with 'wish' y
11. Relative clauses
12. Alternatives of if

During the interview, the participants (47 teachers and 3 ELT experts – assessment and educational measurement experts preferred not to participate in language specific questions as their specialty is not English major.) made clear that the following grammar elements/topics were not represented on the 2021-2022 exam:

1. past perfect continuous
2. *will or going to* used for prediction
3. *will or be going to*
4. present perfect continuous
5. quantifiers *a few, few, a little, little, each none, every, all*
6. *must or can or can't* have + p.p.
7. *would* for past habits
8. compound adjectives
9. narrative tenses
10. future perfect

During the interview, both and teachers (n. 47) and three ELT experts, stated that only eight grammatical elements/ topics were represented on the 2021-2022 exam (%56 coverage/representation of all elements in the curriculum). It is noteworthy that the element/ topic coverage percentage in the 2020-2021 exam was 48%, as in Table (3). Although this can be seen as an unimportant issue (i.e. as student English education within the GCSE is assumed to be grammar-focused), it must be stated that all grammar topics should appear on the exam, from an academic point of view (Syvac, 2018).

Both the teachers and experts (all) aired the view that grammar and vocabulary questions were considerably evident on the 2021-2022 exam in quite intensive but under-representative manner. The teachers (n. 47) expressed that “the grammar questions were a lot”. Actually, they were eleven and seven, respectively, out of a total of forty questions, while on the 2020-2021 exam, they were ten and six, respectively. This is not seen as quite many questions for grammatical elements as not all grammar elements were not represented in both exams. As for the vocabulary elements, the

number is seen as quite little in both, as 79 items were missing, as in Table (3) as 150 were, too, as indicated in Table 4, for both consecutive exams.

A particular issue was raised by teachers during the interview. The teachers (n. 50) and the experts (n. 3 ELT) complained that there were a little number of grammar questions in the exam which, in their own words, “had basics in the M.O.E. curriculum goals but were too advanced as questions for students.” However, it might be argued that teachers should have trained the students on such questions as they knew the exam would embody higher-thinking level questions developed for a student to search for knowledge and widen their horizons through searching for answers beforehand, for the questions do not require memorization-based answers as was before the new method.

It has to be emphasized that recent research mentioned earlier, e.g. Daneshfar et al. (2018), Kamali et al. (2018), and Restaji et al. (2021), proves that DA is a promising area from the assessment of grammar. No teacher or expert dealt with this aspect included in the present research as discussion with them was limited the exam and exam method nature. However, the former studies focused on and investigated the use DA in grammar assessment alone, which is different from the purpose of the present study.

To sum up the results, it is clear that on both the 2020-2021 and the 2021-2022 English exams, all language skills and elements were under-represented, with the complete absence of the representation of both listening and speaking, language functions, and idioms as areas/language elements. However, translation, reading, and grammar as areas received more representation weight than the overall areas of vocabulary and writing.

Therefore, considering data in table (3), Table (4), the representation of those data in Figure (1), Figure (2), and the data in Table (5), the answer to the second set first study question “How far do the 2020-2021 and 2021-2022 GCSE English exams and related exam method represent the language skills and language elements in the M.O.E. Curriculum?” is answered. Result analysis and discussion for data from both the T.E.Q. and the interview have revealed all relating details.

In answer to the second set second study question, “How far does the exam and exam method for the school years 2020-2021 and 2021-2022 represent the novel elements in the two-year curricula?”, it is found that no elements were assessed in both exams. It is noteworthy that there were *no* goals stated for the novel, *Great Expectations*, at the introduction of Haines (2018) nor in the curriculum Chart of Skills in Chappell and Hart (2021a) and Chappell and Hart (2021b), being the M.o.E. official textbooks for the two school years. The teachers (n. 50) and the three ELT experts having the Student’s textbooks, when interviewed, declared that “the M.o.E. did *not* indicate a list of elements for the novel, *Great Expectations*. They said, “the assessment was *poorly* done through merely assessing only the *vocabulary* and *grammar* embedded in the novel, with *no* assessment of understanding such important novel elements as theme, characters, plot, event sequence, setting, climax, end, etc. important to novel learning.”



It is noteworthy that the teachers' views match research addressed earlier (e.g. Burgess, 2022). The teachers explained that "we don't know the reason why those elements weren't covered [represented]." The experts (n. 3 ELT experts) saw that the novel content on the exam was not appropriate. They stated on the questionnaire related to the 2020-2021 exam, "The novel was not properly represented, and this happened to a far extent." They repeated the same statement on the 2021-2022 exam-related questionnaire.

In brief, no elements of 'novel' were addressed on both exams of the two school years. This answers the second set second study question, "How far does the exam and exam method for the school years 2020-2021 and 2021-2022 represent the novel elements in the two-year curricula?"

Conversely, research stresses that all the novel elements, such as theme, characters, plot, event sequence, setting, climax, end, are embedded in any novel content (also, see Thorndike, 1977, for example) and, therefore, they need to be assessed in any novel-incorporating exam.

In answer to the second set third study question, 'What are the strengths and weaknesses of the new assessment method?', first, table (5) establishes a base answer. It reflects a quick teacher and expert evaluation for the new exam method (mirroring the two-year exam design limited to the form of multiple-choice questions marked by a machine) adopted for the GCSE English Exams of the years 2020-2021 and 2021-2022. The information from participants in answer to interview question 4 was obtained, summarized, and analyzed. Table (5) satisfies this purpose by showing how far the method was helpful in the representation of the curriculum language input, as expressed and indicated by both participant teachers and experts.

**Table (5) Teacher and expert views about the new exam method**

	<b>"not helpful"</b>	<b>"moderately helpful"</b>	<b>"much helpful"</b>
<b>Teachers' views</b>	X (n. 41/ 82%)	X (n.9/8%)	
<b>Expert views</b>	X (n. 5/ 100%)		

As indicated in Table (5), the vast majority of teachers and all the experts (n. 41/ 82%) see that the method was not helpful in representing all curriculum language input while a minority see that it was helpful. All the experts (n. 5/ 100%) see that it was not helpful. This suggests that the method was not helpful in assessing all language skills. In the words of assessment and educational measurement psychology expert in answer to question interview no. 6, "A student must write at least six lines in answer to a question on the exam." Two ELT professors see that writing section of the exam must assess all genres in the curriculum, and assessment should not happen at one sitting. It should be ongoing. They said, "a method should allow students to take the exam only when they are ready to take it, and assessment should not happen one time."

The teachers interviewed (n. 41) said, "we are not in favour of the format and the under-represented content of the exam reflected by the method." Both the teachers (n. 47) and the experts (n. 5) interviewed aired the views that the mere dependence on multiple-

choice questions, though appropriate in assessing a wide range of topics and time-saving for administration and marking, which could be the only benefits, caused answers to be much cheated and sometimes guessed by students in such a way that did not exactly reflect the real students' levels in the subject, exactly as observed by the research having two children undergoing the experiences of the GCSE exams. It is noteworthy that these results are in line with research mentioned earlier, e.g. Harris (1969), Rachmat (2019), and Harper et al. (2020: 263).

The experts added that not only were listening and speaking questions not present on the exams, but also the skills of writing and reading were not adequately covered, either. Also, they maintained that even translation as a language element that was in the M.O.E. curriculum did not receive proper nature on the final exam. They stressed, "these are strong weaknesses in the method." "To assess translation, you as an assessor need to ask students to translate; just as when you need assess the four skills, you need to ask students to be involved in their acts," they maintained. They, further, said, "Language functions and idioms were completely missing on the exams while vocabulary did not receive the weight it represented in both curricula." In the words of two ELT experts, "Both exam vocabulary reflected the exceptional use of language and was too advanced for students."

The teachers' and experts' views cope with research findings, discussed earlier, in that students need to be assessed on what they have learned and practised as stated in the goals for the learning they have received, e.g. Spady (1994), Krafft (2012), and Latha and Ravichand, (2019).

It is also found out by the present study that the exams are not committed to the OBE approach of assessment (Spady, 1994) discussed earlier. They were not committed to the goals of the curriculum (i.e. they do not cover the vast majority the skills and language elements number wise and appropriateness wise. The exams did not cater for students' readiness to take them, as mentioned earlier. According to three experts, "Exams should be taken when students are ready to take them." Also, they (n. 5) declared that "the exams did not have continuous shapes." It is known that taking into consideration the assessment of portfolio, demonstrations, performances, monthly tests, quizzes, etc. helps with proper assessment, as discussed earlier. It is noteworthy that much recent research, e.g. Samad (2022: 2) stresses the role of ongoing assessment in language learning.

Thus, the exams either do not fully considerably represent the skills embedded in the M.O.E. Curricula or they are not there altogether, according to the teacher and expert participants. Nor do the exams fully represent the language elements in the curricula (This again answers the second set first study question). They do not assess the elements which exist in any novel (answer to the second set second study question). Two strengths were found for the new method, i.e. coverage of a wide range of study topics and administration and marking easiness. four weaknesses were discovered: inability to assess speaking and the act of writing, disregard of listening, easiness for cheating and guessing by students (answer to the second set third question).

In general, both the teachers (n. 47) and the experts (the three ELT experts) interviewed said the exams reflected by the method covered a considerable amount of translation, reading, and grammar, and they expressed dissatisfaction about the final exams and exam method in general terms as they only contained multiple-choice questions and as they did not cover all the skills and language elements appropriately. In their suppressed anger, they adopted the research findings saying that “the assessment was not actually done on a continuous basis through assignments, quizzes, portfolio, and final exam.” (This acts as an extension to the answer to the second set third study question, ‘What are the strengths and weaknesses of the new assessment method?’).

Although the 2021-2022 Teacher’s Guide (Longman, 2021: 10) confesses that the 2021-2022 curriculum has ongoing assessment elements, it is seen that this statement is superficial and does not happen into practice as students at least rarely attend classrooms in the Third Year Secondary Grade, nor are there any marks obtained during the year for quizzes, portfolio, monthly tests, assignments, etc., except for the those of the final exam. It is, however, well known that, again, ongoing assessment as a method is appropriate for leaning assessment (Samad, 2022: 2). Briefly, the participants’ opinions, backed with a considerable record of research regarding the issues they have raised, were not quite in favour of the exam design and content, nor in favour of the method.

In answer to the second set fourth study question, ‘What are the 2020-2021 and 2021-2022 students’ perceptions about their exams and exam method?’, the following related sub-set questions needed to be addressed:

- 1.1. What are the 2020-2021 students’ perceptions about their exam and exam method?
- 1.2. What are the 2021-2022 students’ perceptions about their exam and exam method?

In answer to question 4.1 above, the S.Q data collected from 2nd year F.o.E. students (n. 100) who attended the 2020-2021 exam were collected and analysed. Table (6) shows these data according to students’ responses to each question in the S. Q. (Appendix 4).

**Table (6) Student responses for the 2020-2021 exam**

Question	Response (1)	Response (2)	Response (3)	
1 <sup>st</sup> Q.	No. (n.97)	Yes (n.3)		
2 <sup>nd</sup> Q.	A number of grammar rules (n. 74)	Listening and speaking (n. 67)	novel	paragraph writing/ essay (n.74)
3 <sup>rd</sup> Q.	To some extent, grammar, but the exam was not related to the curriculum we taught (n.99)			
4 <sup>th</sup> Q.	Multiple-choice (n, 70)	Vocabulary (n. 3)	Inclusion of higher thinking skills (n. 27)	

<b>5<sup>th</sup> Q.</b>	The readings (n.28)	Translation (n. 29)	Two of four alternatives were the same (n.13)	Challenging questions and time insufficiency (n. 30)
<b>6<sup>th</sup> Q.</b>	Yes. Most questions did not address what we had studied (n. 83)	Yes, difficulty to find the right answer among alternatives (n. 55)	Yes, the novel (n.1)	Yes. Difficult vocabulary (n. 21)
<b>7<sup>th</sup> Q.</b>	No. It was only MCQ questions (n.100).			
<b>8<sup>th</sup> Q.</b>	No (n.63)	Yes (n. 37)	N	
<b>9<sup>th</sup> Q.</b>	1. Teacher training 2. Student exam training 3. Curriculum exam representation (n 64)	Improving the reading (n. 26)	Actual translation, not multiple-choice translation (n. 17)	Embedding various types of questions (n. 61)
<b>10<sup>th</sup> Q.</b>	1. Exam must include vocabulary from curriculum 2. More grammar inclusion 3. Attention to improving student levels (n. 55)	Exams must be related to curriculum (n. 83)	Exam should suit all student levels (n.44)	

According to Table (6), GCSE graduates of the year 2020-2021 had various responses per S.Q. question. In answer to the first question (Did the Secondary Stage Exam of English 2020-2021 cover all what you had studied during that year (all types of essays, email writing, surveys, online comments, idioms, language functions, for example, or otherwise?), a vast majority said, “No” (n97). In answer to the second question (What topics were NOT covered?), students views were divided in three categories: reading, translation, and challenging questions and time insufficiency. Student numbers for these views are 28,29 and 30, respectively, while a minority (n 13) said that two of each question four alternatives were the same. It is noteworthy that sometimes, students gave more than one answer. In answer to the third question (Did the exam measure the target language skills as stated in the goals of the Student’s Book and Workbook (listening, speaking, reading, writing, vocabulary, grammar, and translation?), The participants (n. 99) stated “To some extent, grammar, but the exam was not related to the curriculum we taught.” In answer to the fourth question (What do you like about the exam?), students’ views lied mainly between “Multiple-choice” (n, 70), and “Inclusion of higher thinking skills” (n. 27), while a minority mentioned that they liked vocabulary (n.3).

In answer to the fifth question (What do you NOT like about it [the exam]?), participant responses ranged among “The readings” (n.28), “Translation” (n. 29), “Challenging questions and time insufficiency” (n. 30, while a minority mentioned that “two of four alternatives were the same.” (n.13). In answer to the sixth question (Did you face any challenges on the exam? If yes, write some of them?), a clear majority of participants indicated that “Yes. Most questions did not address what we had studied (n. 83) and another majority, sharing more than one response with the first majority, stated that

“Yes, difficulty to find the right answer among alternatives” (n. 55). A minority mentioned that “Yes. Difficult vocabulary” (n. 21).

In answer to the seventh question (Did the exam use various types of questions, such as Fill in the gap questions, paragraph writing, multiple-choice, translation?), all the participants made clear that “No. It was only MCQ questions.” (n.100) In answer to the eighth question (Did the exam trigger all your knowledge, understanding, application, synthesis, analysis, evaluation, and creativity?), a considerable majority stated, “No.” (n.63) while a minority stated, “Yes.” (n. 37) In answer to the ninth question (How do you think future Secondary Stage exam of English can be improved?), a majority states “Teacher training, student exam training, and curriculum-exam representation.” (n 64), another majority sharing a number of views with the former majority, indicated “Embedding various types of questions.” (n. 61), while two minorities mentioned “Improving the reading” (n. 26), and “Actual translation, not multiple-choice translation.” (n. 17) In response to “Feel free to add any suggestions” at the end of the S.Q, a clear majority of participants commented that Exams must be related to curriculum” (n. 83), another majority sharing a number of views with the former majority, believed in “Exam must include vocabulary from curriculum, more grammar inclusion is needed, and attention to improving real student levels” (n. 55), while a considerable minority indicated that “Exam should suit all student levels” (n.44). This answers the second set study question 4.1: What are the 2020-2021 students’ perceptions about their exam and exam method?”

In answer to question 4.2 (What are the 2021-2022 students’ perceptions about their exam and exam method?), the S.Q data collected from 1nd year F.o.M./GCSE-graduate students (n. 50) who attended the 2021-2022 English exam were addressed and tabulated. Table (7) displays students’ responses to each S.Q. question as in Appendix 4.

**Table (7) GCSE graduate responses for the 2021-2022 exam**

Question	Response (1)	Response (2)	Response (3)
1 <sup>st</sup> Q.	No. (n.40)	Yes (n.10)	
2 <sup>nd</sup> Q.	types of writing, essay questions, a number of grammar rules, novel events, and absence of listening and speaking (n. 35)	absence of listening and speaking (n. 15)	
3 <sup>rd</sup> Q.		To some extent, grammar, but the exam was not mostly related to the curriculum (n. 50)	
4 <sup>th</sup> Q.	Question difficulty/challenging and inclusion of higher thinking nature (n. 40)	Question arrangement (n. 10)	
5 <sup>th</sup> Q.	Absence of showing my writing skills and MCQ nature, absence of actual writing (40)	Some challenging questions (n. 10)	
6 <sup>th</sup> Q.	a grammar question related to articles (n. 30)	-	
7 <sup>th</sup> Q.	No. (n. 50). It was only MCQ type		
8 <sup>th</sup> Q.	No (n.14)	Yes (n. 36)	



<b>9<sup>th</sup> Q.</b>	1. Inclusion of essay writing questions 2. coverage what we study 3. distinction of students' levels 4. listening and speaking 5. Embedding various types of questions (n 40)		
<b>Suggestions</b>		Exam must reflect curriculum taught and help to improve students' proficiency levels (n. 30)	Exam should reflect actual student levels (n. 20)

According to Table (7), GCSE graduates of the year 2021-2022 had quite different responses from those of the year 2020-2021. In answer to the first question (Did the Secondary Stage Exam of English 2021-2022 cover all what you had studied during that year (all types of essays, email writing, surveys, online comments, idioms, language functions, for example, or otherwise?), a straightforward majority revealed, "No" (n. 40), while a minority stated, "Yes". In answer to the second question (What topics were NOT covered?), students views showed "types of writing, essay questions, a number of grammar rules, novel events, and absence of listening and speaking" (n. 35) and absence of listening and speaking (n. 15). In answer to the third question (Did the exam measure the target language skills as stated in the goals of the Student's Book and Workbook (listening, speaking, reading, writing, vocabulary, grammar, and translation?), The participants (n. 50) stated "To some extent, grammar, but the exam was not related to the curriculum we taught."

In answer to the fourth question (What do you like about the exam?), students' views lied mainly between "Question difficulty/challenging and inclusion of higher thinking nature" (n. 40), and "Question arrangement" (n. 10). In answer to the fifth question (What do you NOT like about it [the exam]?), participant responses reflected "Absence of showing my writing skills and M.C.Q. nature, absence of actual writing" (n. 40), and "Some challenging questions" (n. 10) In answer to the sixth question (Did you face any challenges on the exam? If yes, write some of them?), a majority of participants indicated that "Yes. a grammar question related to articles" (n. 30).

In answer to the seventh question (Did the exam use various types of questions, such as Fill in the gap questions, paragraph writing, multiple-choice, translation?), all the participants (n.100) made clear that "No. It was only MCQ questions." In answer to the eighth question (Did the exam trigger all your knowledge, understanding, application, synthesis, analysis, evaluation, and creativity?), a considerable majority stated, "yes." (n.36) while a minority stated, "No." (n. 14)

In answer to the ninth question (How do you think future Secondary Stage exam of English can be improved?), a majority stated, "Inclusion of essay writing questions, coverage what we study, distinction of students' levels, and listening and speaking" and "Embedding various types of questions." (n. 40). In response to "Feel free to add any suggestions" at the end of the S.Q, a clear majority of participants commented that

Exam must reflect curriculum taught and help to improve students' proficiency levels (n. 30), and that "Exam should reflect actual student levels" (n. 20).

Of this, the answer to the second set study question 4.2 "What are the 2021-2022 students' perceptions about their exam and exam method?" has been made. Therefore, answer to the second set fourth study question "What are the 2020-2021 and 2021-2022 students' perceptions about their exams and exam method?" has been obtained

Comparing the views of the 2020-2021 student group and those of the 2021-2022 student group, a number of agreements and differences in opinions have been detected. For example, both groups agreed that the exams did not represent all what they had taught, which is comparable to results reached from both the teachers and experts, as shown in Tables (1) and (2), and Figures (1) and (2) and as discussed earlier. Reasons include the fact that students what they have taught and what they were assessed upon. Also, both groups had quite a consensus that the types of writing, essay questions, a number of grammar rules, novel events, and absence of listening and speaking were missing on both exams. Reasons for this, as mentioned earlier, can be referred to the question bank withdrawal rules, which are claimed be the same, disregarding the change in curriculum from Haines (2018) to Chappell and Hart (2021a) and Chappell and Hart (2021b). Again, this result copes with the data obtained from both the teachers and experts analysed and discussed. The only differences were that the teachers and experts indicated more details about areas unincluded in the assessments, such as idioms and language functions, and other novel elements, e.g. characters, theme, climax, etc., as will follow later.

Both groups also generally agreed that grammar was not covered by the curricula, but they added that the exams were not mostly related to the curriculum. Again, this result is line with the results from both the teacher and expert participants. However, the teacher and expert participants added that various types of writing were not included. It is evident that all participants, teachers, experts, and both groups of students, agreed that the exams did not represent the curricula appropriately.

The student groups further had similarities and differences in views regarding what they liked about the exams. Meanwhile the 2020-2021 exam group mainly liked Multiple-choice questions, the other group found it was not a good nature. Nonetheless, both groups liked the question difficulty/challenging question nature and inclusion of higher thinking skills. The only difference was that the 2021-2022 group liked the way the questions were arranged. This is related to the method. However, participants stating this were a minority whose number was ten out fifty total population (20%) and generalisation of verdict cannot be established. The M.o.E. strived for including higher thinking skills, and the majority both groups agree that they were present, which can be another advantage to the exam method, an advantage the teacher and expert participant were not focused on.

From another perspective, while the 2020-2021 group of student participants indicated that they did not like the exam readings, translation, and challenging questions, and time insufficiency, the 2021-2022 reflected that the absence of showing my writing

skills and MCQ nature, absence of actual writing were not welcomed by the group. These differences in views can be attributed to the fact that students of the former group had not been shown an example of the exam while the latter group was quite experienced by the former group's experience. Generally, the results are consistent with the latter group are consistent with those analysed and interpreted from Tables (1) and (2). However, the former groups' finding is different as they are the ones who took the exam, so they could tell time was insufficient, but when it comes to the challenging nature of questions, both groups agree on this aspect of exam, as mentioned previously.

Furthermore, both groups agreed that most questions were not related to the curricula they taught. This is an aspect which both the teachers and experts did not address because of the limited nature the questions they were asked. It has to be asserted that exams addressing what has been taught is inevitable, as proved earlier by literature.

It is also clear that both groups decided that the type of questions related to the assessment method was multiple-choice questions only. This was inferred by what the teachers and expert participants stated earlier.

Both groups had differences regarding whether the exam triggered their knowledge, understanding, application, synthesis, analysis, evaluation, and creativity. The majority of students of the 2020-2021 group indicated that the related exam did not, whereas the majority of 2021-2022 group asserted this happened. This difference refers to the reluctant nature to change. The former students had the first administration of such a new method exam; thus, they felt awed by it, and even were not in a better position to address it as they had had no experience, while the latter group had the advantage of seeing a model: the former group exam and exam method. This was no point the teacher and experts addressed as it was out of focus to them.

Both group views were congruent in suggesting ways for improving the exams and exam method. The majority of both saw that Embedding various types of questions and depending only on curriculum material involved for exams were good interventions. Nevertheless, the 2020-2021 exam group saw that teacher training and student exam training were good options while the 2021-2022 exam group saw that inclusion of essay writing questions, coverage what they studied, distinction of students' levels, listening and speaking inclusion, and embedding various types of questions were the necessary recommendations. However, this difference is meaningless as it is known that these suggestions/ recommendations for improvement supplement one another, as both exams quite share the same characteristics, even appearing in the number of area questions; see Tables (1) and (2). These results largely cope with those reached earlier with respect to teachers' and experts' views regarding 1) necessity of writing type, listening, and speaking inclusion, 2) using other relevant types of questions (as will follow below regarding 'answer to the second set fifth study question'.) GCSE graduates gave different view from those teachers and experts: the former exam group mentioned teacher training and student exam training, and the latter group mentioned the idea of exam prerequisite to have the quality to distinguish among students' levels, which is considered an advantage for the students' views over both the teachers' and experts', as both did not mention these important remarks.

As for answer to the second set fifth study question, ‘What are particular plans for improvement, if needed?’, the vast majority of the teachers (n.47) and experts (n.5) interviewed stated that there were other types of questions which could be used for better assessment of novel elements, such as essay or paragraph writing questions as they are a bit more difficult to cheat and can provide a rather accurate assessment of certain language skills, such as writing; and language elements, such as translation. They said that error-correction questions were useful as well, particularly in grammar and maintained that dialogues and situation questions are suitable to assess language function as was embedded in the old method. They also suggested the coverage of most language skills and elements on the exam, and the inclusion of both listening and speaking on the exam, providing listening is played by a knowledgeable teacher, and providing speaking is recorded. They added that writing, as well as translation, needs to be practised on the exam.

Even, the GCSE graduates in the two groups showed even stronger, quite more comprehensive views. The majority of them saw that exam suitability of all student levels, teacher and student training on the exam, inclusion of more grammar and vocabulary from the curriculum, exam design to care about raising students’ levels (2021-2021 group) are their suggestions for plans of action. Also, the GCSE graduates (2021-2022) aired the views that exam must reflect curriculum taught and help to improve students’ proficiency levels; they should cater for the inclusion of essay writing questions, distinction of students’ levels, listening and speaking inclusion. Both GCSE graduate groups shares the suggestions of including various types of questions and representation of curriculum on the exam.

This answers second set fifth study question. In fact, all the teachers’ and experts’, and students’ recommendations/suggestions are considerably similar to research evidence which has been discussed earlier, e.g. Islam et al. (2021), Sychev et al. (2020), Frost (2005), Arif et al. (2021), Sychev et al. (2020), Harper et al. (2020: 263), Latha and Ravichand (2019:1), Vu et al. (2022: 41), Echcharfy (2019: 361), Popovic (2020a). It is even surprising that students’ recommendations were quite a little more comprehensive than the other participants.

## CONCLUSIONS AND RECOMMENDATIONS

Based on actual data collected from the field stakeholders (teachers and students) and academic experts and later analysed, six hazards have emerged and need quick action just as the new assessment method is still into effect. Below is a record of those hazards (extracted from the nature and characteristics of the GCSE English exam and exam method) and scrutinization of strategies for intervention.

First, no performance-based assessment of the three skills of listening, speaking, and writing was included in both the 2020-2021 and 2021-2022 GCSE English exams. Knowledge about writing, not writing was assessed. Besides, most of reading types and genres were absent on the exams. This needs an invention action to *include* all/ most of the missing areas, skills, genres, and types of reading, and find ways to monitor the students’ performance appropriately and fairly, i.e. by providing knowledgeable

invigilators for playing listening audios, as well as introducing into the system interviewees, i.e. specialized trusted teacher experts for speaking assessment, and making available equipment necessary for listening assessment, such as stereo speakers, computers, laptops, for example, in exam venues. It is advised that all the skills should be assessed in ballance with one another.

Second, the sole type of multiple-choice questions was used for assessment. One type cannot be said to enough. It is well known that in order to assess performance skills, a teacher needs to ask students to *perform/do* the task which demonstrates the performance levels/ abilities of those students. To assess writing, students need to write; to assess listening, students need to listen; to assess speaking, students need to speak. Carrying out those tasks out (as well as reading, of course) by students with clear instructions and procedures on the part of teachers/invigilators will inform of real students' abilities.

Third, the exams did not cover all the grammatical elements in the curricula. It is strongly advisable that all/ the vast majority of the grammatical structures in the curriculum need to exist on the exam. Grammar is extremely important to all the four macro-language skills/areas and language elements with respect to structuring and understanding the language in communication (Kosimov, 2022).

A comprehensive exam grammar coverage is then much suggested to enable students to show how far they have mastered this controlling, major language feature/element. DA is suggested as an assessment strategy. Also, a Determine-the-Incorrect-Part-of-a-Sentence/an error-correction task is much advisable together with the multiple-choice type questions for pen and paper/ electronic grammar assessment.

Fourth, translation was not carried out by students in the exams. It has to be made clear that translation needs to be *performed* by students when it is assessed (Alenezi, 2020). Students need to be involved in the act of translating texts, not merely choosing/guessing the correct translation out of four alternatives. Choosing/guessing the correct answer does not tell if a student is truly accurately capable of knowing the meaning of terms/vocabulary imbedded in the target text, not does it tell the students are necessarily able to structure a sentence well.

Fifth, Language functions, vocabulary, and idioms were not represented in such an appropriate weight as much as they were represented in the language input of both year curricula. All language functions have to exist in teaching the language (Nugroho and Rekha, 2020: 2). Teaching, and hence assessing, language functions is needed for learners to meet the conditions of communication competence in an intercultural society (Echcharfy (2019: 361). They should be assessed through dialogue questions as well (Jennach et al., 2021). Then, all language functions need to be well addressed in terms of variation. Even when one argues that most of them were taught in the previous two years of the Secondary Stage, it has to be clarified that they still need to appear as reminders not only in GCSE curriculum but also on the final stage exam as well. It is also known that vocabulary and idioms raise the competency of the learner (Sari et al., 2021); therefore, they need to be well represented on the stage exam, too.



Finally, the exam in general does not follow the OBE principles, as stated by Spady (1994), as explained earlier. It does reflect the goals embedded in (or inferred from) the Teacher's Guide. Students' appropriate circumstances to be tested were not taken into consideration, as the exam is only taken in one day and does not cater for students' readiness to take the exam. According to the words on one academic expert interviewed, "Students should only take the exam when they are ready to take it, and the date and time should not be imposed on them." This practice of a fixed exam date needs to be changed to encourage better student abilities. The idea of one-shot exam is not appropriate for perfect assessment as assessment should be continuous, as informed earlier. Also, teachers and students should practice mimic assessments beforehand. The assessment/exam has to cope with OBE principles and it has to be performance-based.

In conclusion, six negative assessment practices are found on final assessment of students' level of English: absence of performance tasks for three major language skills of listening, speaking, and writing, and most genres and skills were about knowledge of writing. Reading was under-represented though performed on exams. Malpractices, such as dependence on one type of questions, disregard of most grammar topics, vocabulary items, all idioms, and all language functions in curriculum, unpractised student translation, and in compliance with OBD values well all evident whether according to the stakeholders or according to research. Reverse actions are suggested/recommended in solving these problems. These include performance-based treatment and sub-skill coverage showing students' abilities of the four macro-skills, performance of the listening, speaking and writing skills as well as the translation element, coverage of grammatical points, vocabulary, idioms, and language functions in curriculum, and enumerating language functions necessary for life, varying question types according to language skill/element, training teachers and students on exam-related matters, and commitment to the OBE approach for assessment.

### Directions for further research

As this study was only conducted to teachers and students as stakeholders, it is preferred that parents' views, too, should be investigated. The following themes can also be suggested for research purposes:

1. Evaluating other subject exams of GCSE students
2. Assessing the social impact of the new assessment method
3. Assessing the skill performance of GCSE students
4. Case- studying of GCSE graduates to assess their problem solving and critical thinking skills
5. Investigating the grounds of cheating on college-level exams
6. Determining the time when multiple-choice questions can truly and accurately assess students' language skills
7. Finding alternative possible methods for GCSE assessment

### References

- Abduh, M. Y. M. (2021). Full-Time Online Assessment during COVID-19 Lockdown: EFL Teachers' Perceptions. *Asian EFL Journal*, 28(11), 26–46.  
[https://www.researchgate.net/publication/349103573\\_Full-](https://www.researchgate.net/publication/349103573_Full-)

- time\_Online\_Assessment\_during\_COVID\_19\_Lockdown\_EFL\_Teachers%27\_Perceptions
- Ahmed, A. (2019). Training needs of Secondary School teachers in light of the requirements of the new educational system in Egypt, a field study, *Journal of the Banha Faculty of Education*, 3(119 July 3c), 453-530. DOI: 10.21608/jfeb.2019.61338
- Akongoh, R. N. (2021). Teacher-Based Assessment of Speaking in Cameroonian Secondary Schools: The Impact of Teacher Training. *Journal of English Language Teaching and Applied Linguistics*, 3(2), 01–11. <https://doi.org/10.32996/jeltal.2021.3.2.1>
- Alenezi, A. M. (2020). Task-Based Approach in Teaching Translation: A Case Study in Jouf University. *Higher Education Studies*, 10(2), 189-196. <https://doi.org/10.5539/hes.v10n2p189>
- Burgess, A. (2022). Novel Literature. <https://www.britannica.com/art/novel/Propaganda>
- Chappell, P. & Hart, C. (2021a). New Hello Term 1 Student's Book and Workbook. York Press Ltd.
- Chappell, P. & Hart, C. (2021b). New Hello Term 2 Student's Book and Workbook. York Press Ltd.
- Daneshfar, S., Aliasin, S. H., & Hashemi, A. (2018). The Effect of Dynamic Assessment on Grammar Achievement of Iranian Third Grade Secondary School EFL Learners. *Theory & Practice in Language Studies*, 8(3). DOI: <http://dx.doi.org/10.17507/tpls.0803.04>
- Estaji, M., & Ameri, A. F. (2020). Dynamic assessment and its impact on pre-intermediate and high-intermediate EFL learners' grammar achievement. *Cogent Education*, 7(1), 1740040. <https://doi.org/10.1080/2331186X.2020.1740040>
- Ewiss, M. (2021). Assessment of School's Evaluation Systems in Egypt Before and During COVID, *American, Journal of Humanities and Social Sciences* 5(3):7-22. <https://www.researchgate.net/publication/349692059>
- Fauzi, I. (2018). The Effectiveness of Skimming and Scanning Strategies in Improving Comprehension and Reading Speed Rates to Students of English Study Programme. *Register Journal*, 11(1), 101-120. DOI: 10.18326/rjt.v11i1.101-120
- Frost, R. (2005). The question types. British Council. Fulcher, G., & Davidson, F. (2007). Language testing and assessment: An advanced resource book. Routledge.
- Harper, R., Bretag, T., & Rundle, K. (2021). Detecting contract cheating: examining the role of assessment type. *Higher Education Research & Development*, 40(2), 263-278. <https://doi.org/10.1080/07294360.2020.1724899>
- Harris, D. P. (1969). Testing English as a Second Language. McGraw-Hill.
- Ibrahim, W., & Othman, Z. (2021). Error Analysis on the Malaysian Students' Writing for ESL Classroom. *LSP International Journal*, 8(2), 55-65. <https://doi.org/10.11113/lspi.v8.17936>
- Islam, M.S., Hasan, M.K., Sultana, S. (2021). English language assessment in Bangladesh today: principles, practices, and problems, *Language Testing in Asia*, 1(21), 1-21. <https://doi.org/10.1186/s40468-020-00116-z>

- Jannah, R., Hidayat, D. N., Husna, N., & Khasbani, I. (2021). An item analysis on multiple-choice questions: a case of a junior high school English try-out test in Indonesia. *Leksika: Jurnal Bahasa, Sastra dan Pengajarannya*, 15(1), 9-17. DOI: 10.30595/lks.v15i1.8768
- Javid, L. (2014). The comparison between Multiple-Choice (MC) and Multiple True-false (MTF) Test Formats in Iranian Intermediate EFL learners' vocabulary learning. *Procedia - Social and Behavioral Sciences*, 98, 784–788. <https://doi.org/10.1016/j.sbspro.2014.03.482>
- Jannach, D., Manzoor, A., Cai, W., & Chen, L. (2021). A survey on conversational recommender systems. *ACM Computing Surveys (CSUR)*, 54(5), 1-36. <https://doi.org/10.1145/3453154>
- Kamali, M., Abbasi, M., & Sadighi, F. (2018). The effect of dynamic assessment on L2 grammar acquisition by Iranian EFL learners. *International Journal of Education and Literacy Studies*, 6(1), 72-78. <http://dx.doi.org/10.7575/aiac.ijels.v.6n.1p.72>
- Khansir, A. A., & Pakdel, F. (2018). Place of error correction in English language teaching. *Educational Process: International Journal*, 7(3), 189. <http://dx.doi.org/10.22521/edupij.2018.73.3>
- Kitjaroonchai, N. & Kitjaroonchai, T. & Sanitchai, P. (2022). The Effects of Process Genre-Based Writing and Process Writing Approaches on Asian EFL Pre-University Students' Writing Performance. *Journal of Language Teaching and Research*. DOI: 10.17507/jltr.1304.19
- Kosimov, A. (2022). THE IMPORTANCE OF CHOOSING APPROPRIATE ASSESSMENT TOOLS IN LANGUAGE TEACHING AND ITS IMPACT TO SECOND LANGUAGE ACQUISITION. *British View*, 7(1).
- Krafft, C. (2012). Challenges facing the Egyptian education system: Access, quality, and inequality," Survey of Young People in Egypt Policy Brief no. 2. Cairo: Population Council, (2012). Available at <https://core.ac.uk/download/pdf/229798627.pdf/>
- Latha, S. P., & Ravichand, M. (2019). Role of Language Testing and Evaluating in Outcome Based Education. [https://www.researchgate.net/profile/Ravichand-Mandalapu/publication/361761697\\_Role\\_of\\_Language\\_Testing\\_and\\_Evaluating\\_in\\_Outcome-Based\\_Education/links/62c3f58f412e4c2aaeab314b/Role-of-Language-Testing-and-Evaluating-in-Outcome-Based-Education.pdf](https://www.researchgate.net/profile/Ravichand-Mandalapu/publication/361761697_Role_of_Language_Testing_and_Evaluating_in_Outcome-Based_Education/links/62c3f58f412e4c2aaeab314b/Role-of-Language-Testing-and-Evaluating-in-Outcome-Based-Education.pdf)
- Longman (2021). Teacher's Guide.
- Ma'arif, A., Abdullah, F., Fatimah, A., & Hidayati, A. (2021). Portfolio-based assessment in English language learning: Highlighting the students' perceptions. *J-SHMIC: Journal of English for Academic*, 8(1), 1-11. DOI: [https://doi.org/10.25299/jshmic.2021.vol8\(1\).6327](https://doi.org/10.25299/jshmic.2021.vol8(1).6327)
- Miralpeix, I., & Muñoz, C. (2018). Receptive vocabulary size and its relationship to EFL language skills. *International Review of Applied Linguistics in Language Teaching*, 56(1), 1-24. <https://doi.org/10.1515/iral-2017-0016>
- Nugroho, A., & Rekha, A. (2020). Speech acts of requests: A case of Indonesian EFL learners. *Journal of English Language Teaching and Linguistics*, 5(1), 1-16.
- Nurwahidah, N., Hamidi, M. R., Husna, N., Farkhan, M., Hidayat, D. N., & Alek, A. (2022). English Teachers' Challenges in Constructing Remote English Summative Assessment during Covid-19 Pandemic. *Linguistic, English*

- 
- Education and Art (LEEA) Journal*, 5(2), 183-193.  
DOI: <https://doi.org/10.31539/leea.v5i2.3534>
- OECD (2015). SCHOOLS FOR SKILLS: A NEW LEARNING AGENDA FOR EGYPT, OECD Report (2015). Available at <https://www.oecd.org/countries/egypt/Schools-for-skills-a-new-learning-agenda-for-Egypt.pdf>
- O'Loughlin, K. (2006). Learning about second language assessment: Insights from a postgraduate student on-line subject forum. *University of Sydney Papers in TESOL*, 1, 71-85.
- Oxford Business Group (2018). Reform of Egypt's public educational curriculum begins. <https://oxfordbusinessgroup.com/overview/system-overhaul-extensive-reform-curriculum-enters-itsinitial-phase>
- Popovic, M. (2020a). Informative manual evaluation of machine translation output. In Proceedings of the 28th International Conference on Computational Linguistics, pages 5059–5069, Barcelona, Spain (Online). International Committee on Computational Linguistics. DOI:10.18653/v1/2020.coling-main.444
- Popovic, M. (2020b). Relations between comprehensibility and adequacy errors in machine translation output. In Proceedings of the 24th Conference on Computational Natural Language Learning, pages 256–264, Online. Association for Computational Linguistics. DOI:10.18653/v1/2020.conll-1.19
- Putri, D. & Susanti, A. (2022). EDMODO AS ONLINE ASSESSMENT TOOL: TO IMPROVE EFL STUDENTS' READING COMPREHENSION, *Jurnal Satya Widya*, 38 (1), 12 – 24. DOI: <https://doi.org/10.24246/j.sw.2022.v38.i1.p12-24>
- Rachmat, N. A. (2019). *Senior High School English Teachers' Perceptions in Using Multiple-Choice Question as a Summative Assessment* (Doctoral dissertation, Universitas Muhammadiyah Yogyakarta).
- Robiasih, H., & Lestari, T. (2020). Formative Assessment Performed by High School Teachers in the Pandemic Era. *Loquen: English Studies Journal*, 13(2), 80-87. DOI: <http://dx.doi.org/10.32678/loquen.v13i02>
- Saavedra, J. (2019). Shaking up Egypt's public education system, Education for Global Development, (2019). Available at <https://blogs.worldbank.org/education/shaking-egypts-public-education-system/>
- Samad, M. A. (2022). Understanding EFL Learners' Errors in Language Knowledge in Ongoing Assessments. *English Language Teaching*, 15(9), 1-8. <https://doi.org/10.5539/elt.v15n9p1>
- Sari, S. N., & Aminatun, D. (2021). STUDENTS' PERCEPTION ON THE USE OF ENGLISH MOVIES TO IMPROVE VOCABULARY MASTERY. *Journal of English Language Teaching and Learning*, 2(1), 16-22. DOI: <https://doi.org/10.33365/jeltl.v2i1.757>
- Seow, A. (2002). The Writing Process and Process Writing. DOI: 10.1017/CBO9780511667190.044
- Spady, W. G. 1994, "Choosing outcomes of significance", *Educational Leadership*, 51, 6, 18-22. In Latha, S. P., & Ravichand, M. Role of Language Testing and Evaluating in Outcome Based Education.
- Srinivasan, P. (2016). EVALUATION IN EDUCATION- EDUCATIONAL EVALUATION

- Sychev, O., Anikin, A., & Prokudin, A. (2020). Automatic grading and hinting in open-ended text questions. *Cognitive Systems Research*, 59, 264-272. <https://doi.org/10.1016/j.cogsys.2019.09.025>
- Syvak, O. (2018). Importance of Grammar in ESP. *Science and Education a New Dimension*. VI (183), DOI: 10.31174/SEND-Ph2018-183VI54-13
- Thorndyke, P. W. (1977). Cognitive structures in comprehension and memory of narrative discourse. *Cognitive psychology*, 9(1), 77-110.
- Weimer, M. (2018). Multiple-choice tests: Revisiting the pros and cons. Retrieved from *Faculty Focus*: <https://www.facultyfocus.com/articles/teaching-professor-blog/multiple-choice-tests-pros-cons>.
- Vu, D., Lian, A., Siriyothin, P. (2022). Integrating Natural Language Processing and Multimedia Databases in CALL Software: Development and Evaluation of an ICALL Application for EFL listening Comprehension, *Computer Assisted Language Learning Electronic Journal (CALL-EJ)*, 23(3), 41-69.

حسن، ع. ا. م.، عبد الله محمد، الانصاري، محمد صبرى، عبد الرازق، و صابرين نشأت. (2021). الاحتياجات التدريسية لمديري المدارس الثانوية العامة في ضوء نظام التعليم الجديد في مصر، *مجلة العلوم التربوية- كلية التربية بقنا*، 49(49)، 14-264. DOI: 10.21608/maeq.2021.98589.1053