Tabaran Institute of Higher Education
International Journal of Language Testing
Vol. 15, No. 1; 2025, March

ISSN 2476-5880
https://www.ijlt.ir

Examining the Elicited Imitation Test in an EFL Classroom: Insights from Language Assessment and Student Perception

Ghadah Albarqi[1]*

ARTICLE INFO

Article History:
Received: April 2024
Accepted: May 2024

KEYWORDS
Elicited imitation
Listening
Oxford Placement Test
Speaking
Study success

ABSTRACT

The Elicited Imitation Test (EIT) is widely recognized for its reliability in research settings as a proficiency assessment tool. However, there exists a need to examine its predictive validity in English as a Foreign Language (EFL) classrooms. This study investigates the extent to which the EIT, alongside the Oxford Placement Test (OPT), can predict students' academic achievements in an English for Academic Purposes course, including overall grade point average and scores in listening, speaking, grammar, and vocabulary. The study also examines the relationship between students' perceptions of their listening and speaking skills and their EIT performance. The study involves 41 participants, with data analysis conducted using both regression and correlation methods. Results show that the EIT significantly predicts students' grade point average and language skills. Students' self-perceived speaking and listening abilities reasonably align with their actual performance on the EIT, and it seems that factors related to comprehension weigh heavily in their considerations. These findings have significant implications for EFL research and pedagogy.

1. Introduction

The Elicited Imitation Test (EIT) has attracted considerable interest in research settings as a means of assessing second language (L2) oral proficiency through precise sentence repetition (see Akbary et al., 2023; Kim et al., 2024; Ortega et al., 2002; Solon & Park, 2024; Suzuki & DeKeyser, 2015; Wu et al., 2021; Wu & Ortega, 2013). In the EIT, participants are required to repeat sentences ranging from 7 to 19 syllables, with an emphasis on accurate repetition (see Appendix B). Research findings strongly suggest that EIT serves as a valuable tool for evaluating processing efficiency, particularly in terms of automaticity, which pertains to the ease with which individuals comprehend and produce spoken language (see Gaillard & Tremblay, 2016; Van Moere, 2012).

Researchers often opt for the use of EIT due to its affordability and practicality. This assessment offers a cost-effective and resource-friendly means of measuring language proficiency, making it highly conducive to research endeavors (Erlam, 2006; Gaillard & Tremblay, 2016; Kim et al., 2016). Another notable advantage of EIT is its potential for facilitating comparisons of research results across various second languages. Researchers can use the EIT as a common benchmark, allowing for a degree of standardization in assessing language proficiency (Kim et al., 2016). Nonetheless, there are certain gaps in this research area that require further investigation. Firstly, it is important to examine the predictive validity of EIT within novel contexts such as English as a Foreign Language (EFL) classrooms (Gómez-Benito et al., 2018; Isbell & Son, 2022; Zumbo, 2007). This effort has the potential to significantly enhance the EIT's applicability, moving it from research settings into practical real-world use. Moreover, there is a need to validate the assumption that within the EFL context, learners' listening ability is strongly aligned with their EIT performance (Wu et al., 2021). Validating this assumption is important for gaining insights into the multifaceted nature of language assessment, where learners' self-

[1] Foreign Language Department, Taif University, Taif, Saudi Arabia. Email: gadah.g@tu.edu.sa

Cite this paper as: Albarqi, G. (2025). Examining the elicited imitation test in an EFL classroom: Insights from language assessment and student perception. *International Journal of Language Testing*, *15*(1), 20–39. https://doi.org/10.22034/ijlt.2024.452903.1339

perceived abilities can function as valuable assessment resources. Finally, researchers need to explore participants' perceptions of the challenges tied to comprehending and producing EIT sentences. This investigation can provide insights into whether the challenges encountered in EIT performance are rooted in comprehension or production skills within EFL contexts.

2. Review of Literature

2.1. EIT Research and Validation

The EIT has been used in L2 studies to evaluate learners' oral linguistic proficiency (Ellis, 2005; Erlam, 2006; Gaillard & Tremblay, 2016; Ortega et al., 2002; Wu & Ortega, 2013). Researchers have explained the EIT's processing, highlighting that the process begins with the reception of sentences through the auditory system (Bley-Vroman & Chaudron, 1994). These sentences are then transformed into a mental representation through the comprehension system, temporarily stored in short-term memory, and articulated through the production system (Bley-Vroman & Chaudron, 1994). This perspective suggests that the EIT involves a reconstructive process rather than mere rote repetition. Its effectiveness lies in its ability to elicit the imitation of spoken sentences, thereby accessing learners' implicit grammatical knowledge, which can be applied in real-time situations (Gass, 2018; Wu & Ortega, 2013).

Various validation methods have been employed within research contexts. External validity has been evaluated through the comparison of EIT scores with independently collected measures of language proficiency. These comparisons have consistently demonstrated strong alignment with various standardized tests (see Kim et al., 2016; Kostromitina & Plonsky, 2021). Furthermore, researchers have explored the correlations between EIT scores and established measures of oral language proficiency encompassing Complexity, Accuracy, and Fluency (CAF) dimensions, which have been extensively validated in L2 research (e.g., Kim et al., 2016; Park et al., 2020; Skehan et al., 2016). L2 studies have unveiled meaningful associations between EIT and specific CAF dimensions. For instance, noteworthy connections have been observed between EIT scores and speech rate (Kim et al., 2016; Tracy-Ventura et al., 2014; Wu & Ortega, 2013), accuracy (Albarqi, 2024; Kim et al., 2016; Park et al., 2020), and lexical diversity (Park et al., 2020; Wu & Ortega, 2013). These findings suggest that an individual's performance on the EIT is linked to their overall linguistic ability.

Another approach to validating the external validity of EITs involves utilizing self-assessments of language proficiency. While this method has been employed in a limited number of recent studies, it has yielded insightful results. For instance, Bowden (2016) observed significant correlations between individuals' self-assessments of their speaking and listening skills and their EIT scores. In a more recent investigation conducted by Wu et al. (2021), researchers employed a self-diagnostic survey to explore how participants perceived the factors influencing their performance on the EIT. The results of this study revealed that participants' comprehension ability significantly predicted their performance on the EIT. Moreover, the study indicated that listening ability was a strong predictor of participants' EIT performances. It has been argued that in EFL settings, it is quite common for learners to demonstrate better skills in listening than in speaking (Wu et al., 2021). Nevertheless, additional research is required to confirm this assumption in diverse contexts, including EFL classrooms.

While existing validation research strongly supports the use of EIT scores as indicators of general oral proficiency for research purposes, it has been argued that determining the validity of EIT scores cannot be simplified to just looking at a reliability score and how it correlates with another measure as a reference point (Isbell & Son, 2022; Kane, 2013; Révész & Brunfaut, 2021). To advance this research further, Isbell and Son (2022) built upon Kane's (2013) validity framework. Kane's (2013) framework proposes that we can establish the validity of a test by analyzing how we interpret and use the test scores. This involves tracing "a network of inferences and assumptions leading from the test performances to the conclusions to be drawn" (Kane, 2013, p. 8). Within this framework, a pertinent facet of validity involves the concept of explanation. In essence, this means that the results of the test should align with what individuals understand or can accomplish in the subject being tested. In their study, Isbell and Son (2022) came to the conclusion that a gap still exists in this aspect. Specifically, EIT scores need to demonstrate the ability to account for general oral proficiency across diverse learner subgroups, such as EFL learners (see also Gómez-Benito et al., 2018). In summary, there is a need to assess the degree to which EIT scores can predict the academic performance of EFL students in their

studies. This endeavor can enhance the validity of the EIT beyond research settings and into pedagogical applications.

Furthermore, while the EIT is believed to assess a learner's L2 linguistic competence by capturing their implicit knowledge as they comprehend and reproduce the sentences (Davis & Norris, 2021; Van Moere, 2012), it can be challenging to determine whether the difficulty in sentence repetition arises from comprehension or production issues when failure occurs (Hood & Lightbown, 1978; Vinther, 2002). This ambiguity has led some researchers to question its acceptability (see Vinther, 2002). In addition, a gap exists in our knowledge regarding how participants perceive the challenges inherent in comprehending and reproducing EIT sentences. Few researchers attempted to delve into the factors influencing comprehension and production of the EIT using a self-diagnostic survey that highlights participants' perceptions of these factors (see Wu et al., 2021). Their findings reveal that factors related to comprehension hold a more prominent place in the minds of EFL learners. This line of research is crucial as it enriches our understanding of the challenges learners confront during both the comprehension and production phases of the EIT.

2.2. Academic Success in EFL Classroom

Academic success in an EFL classroom is a major concern for students and institutions. It typically refers to students' achievement in language proficiency, comprehension, and overall performance. Predicting grade outcomes can assist instructors and curriculum designers in devising interventions aimed at enhancing learner performance, and offering additional support to those who may require it to elevate their levels of achievement (Daller & Yixin, 2017).

L2 literature indicates that predictive research has predominantly centered on predicting L2 success for international students who study abroad (see Daller & Yixin, 2017; Feast, 2002). Researchers have traditionally employed internationally recognized standardized English tests like TOEFL and IELTS. These tests have long been considered fundamental indicators of one's proficiency in a foreign language (Daller & Yixin, 2017; Yu, 2014). The rationale behind this approach lies in the belief that a strong command of the target language is a pivotal predictor of achievement in L2 learning (Daller & Phelan, 2013; Daller & Yixin, 2017; Woodrow, 2006). Such standardized tests are designed to evaluate an individual's language proficiency comprehensively, encompassing various language skills such as listening, reading, writing, and speaking. Consequently, they are considered reliable indicators of learners' linguistic abilities, which are highly relevant to their academic performance in L2 contexts (Feast, 2002; Wongtrirat, 2010; Woodrow, 2006).

However, research showed that standardized tests indicated weak to moderate predictive ability on study success. For example, IELTS indicated weak to moderate correlations with students' GPA (see Woodrow, 2006), and TOEFL showed weak predictive ability on GPA (Wongtrirat, 2010). Researchers have raised concerns about these tests (see Ariamanesh et al., 2023; Daller & Yixin, 2017; Souzandehfar, 2024), emphasizing that the specialized preparation for tests like IELTS transforms them from assessments of English proficiency into evaluations of students' test-taking abilities (see Daller & Yixin, 2017; Yu, 2014). Hence, an alternative approach to language assessment has been proposed, advocating for the use of more cost-effective and easily administered tests like the Oxford Placement Test and C-tests (see Daller & Phelan, 2013; Daller & Yixin, 2013; Doró, 2011; Dörnyei & Katona, 1992). Both tests have shown strong predictive validity in anticipating student academic success. However, the OPT holds an advantage as it aligns its scores with the Common European Framework of Reference for Languages (CEFR), making it particularly appealing to L2 researchers (see Doró, 2011; Duran-Karaoz & Tavakoli, 2020).

Within the context of the EFL classroom, researchers also have examined other predictors of EFL students' success such as motivation, anxiety, and learning style (see Dörnyei & Chan, 2013; Kim & Kim, 2011). These studies provide valuable insights into a student's preparedness for EFL courses. However, there remains an unexplored avenue within EFL classrooms, specifically using oral modality assessment such as the EIT, to predict students' achievement. The EIT offers a unique perspective by being conducted in the oral modality. Unlike traditional written assessments, EIT is believed to tap into receptive and productive oral skills as well as implicit language knowledge (Gass, 2018). Likewise, EIT performance provides valuable insights into learners' command of vocabulary and grammar, as well as their ability to use them fluently and with reasonably clear pronunciation (Wu & Ortega, 2013). It has

been argued that, while vocabulary, grammar, pronunciation, and fluency are not standalone measures of overall communicative oral proficiency, they significantly influence speaking test scores (Wu & Ortega, 2013). This implies that the EIT can be a valuable tool for assessing various aspects of language competency in the EFL context, offering a more comprehensive view of learners' abilities beyond traditional written assessments.

3. Aims of the Study

The principal aim of this study is to investigate the predictive validity of the EIT within a novel context, particularly focusing on its applicability in an EFL classroom. This exploration aims to bridge the gap between research settings and practical real-world usage of the EIT. The study also aims to assess its ability to predict EFL students' achievement alongside another independent test (the OPT), thereby strengthening the external validity of the assessment. Secondly, the study aims to validate the assumption that within the EFL context, learners' listening ability closely aligns with their EIT performance (Wu et al., 2021). Finally, the study also endeavors to delve into participants' perceptions of the challenges associated with comprehending and reproducing EIT sentences, offering insights into whether these challenges predominantly stem from comprehension or production skills within an EFL context.

RQ1. To what extent can EIT and OPT predict L2 learners' L2 achievement in an EFL classroom in terms of GPA, listening, speaking, grammar, and vocabulary?
RQ2. How well do students' perceptions of their listening and speaking abilities align with their EIT performance?
RQ3. To what extent do students' EIT performance align with their perception of their EIT performance and factors that affect their comprehension and production of EIT sentences?

4. Method

4.1. Participants and Settings

In this study, a total of 44 Arabic L1 speaking students initially took part, but three students did not complete all the required tasks and were subsequently excluded. The students, aged between 18 and 21 (M = 18.87, SD = .93), were enrolled in their first year at the Pharmacology College at a state university in Saudi Arabia. Before starting college, they had undergone seven years of English language education, throughout their primary, intermediate, and secondary schools. These students were enrolled in an Intensive English for Academic Purposes course (12 hours per week). Each participant provided written consent, demonstrating their understanding of the research's ethical principles and their willingness to participate. Furthermore, participants were assured that their test results would be handled with utmost confidentiality, and the use of their academic records would be carried out ethically.

4.2. Instruments

4.2.1. **Oxford Placement Test.** In this study, the grammar and vocabulary sections of the OPT were used to assess the linguistic knowledge of L2 learners (Allan, 2004). The OPT consisted of 60 questions, with 1 point awarded for each correct answer. The internal consistency of the OPT was assessed using Cronbach's alpha, yielding an acceptable reliability coefficient of $\alpha = .76$. This test was chosen because it is a written test that places emphasis on grammar and vocabulary whereas the EIT predominantly evaluates oral language skills, specifically listening and speaking. According to Duran-Karaoz and Tavakoli (2020), the OPT is more likely to assess participants' declarative knowledge while the EIT is more likely to assess their procedural knowledge. Moreover, the use of the OPT offers the advantage of score alignment with the CEFR (see Appendix A). By incorporating both tests into the research design, a comprehensive exploration can be conducted to determine the extent to which each test exhibits predictive validity across diverse language modalities.

4.2.2. **Elicited Imitation Test.** This study employed the EIT, which was developed by Wu and Ortega (2013). The EIT comprises 30 sentences, characterized by an increase in syllable count from 7 to 19. Participants were provided with a single listening opportunity for the sentences, followed by a 3-second pause before their repetition. Scoring of the sentences was executed on a scale ranging from 0

to 4 points (see Appendix C). The EIT's maximum score is 120, where a perfect repetition warrants 4 points. Repetitions with form or content changes receive 2 points, half repetitions or less get 1 point, and single-word repetitions or inability to repeat receive a score of zero.

To ensure the reliability of the scoring process, the author and a second rater independently coded 10% of the dataset. High interrater reliability was achieved, with a coefficient of .96. Any disagreements were resolved through discussion. Following this assessment, the first rater proceeded to code the remaining dataset. The internal consistency of the EIT was assessed using Cronbach's alpha, yielding a robust reliability coefficient of $\alpha = .95$. This result closely aligns with the findings reported by Wu and Ortega (2013) and Tracy-Ventura et al. (2014).

4.2.3. **Academic Performance.** Participants underwent a 12-hour per week Intensive English for Academic Purposes (IEAP) course. This intensive course was designed to cultivate proficiency in various language skills, encompassing listening, speaking, grammar, vocabulary, reading, and writing. The cumulative score for these skills was 90. Therefore, a total score of 90 was used as the GPA score. The use of GPA in educational research is strongly advocated by researchers (e.g., Bacon & Bean, 2006) due to its strong correlation with various other variables, such as motivation, achievement scores, and teamwork.

The language assessments were commonly prepared and administered by the testing unit at the English Language Centre at the end of the semester. Each language skill was evaluated on a scale of 15 points. For grammar, vocabulary, and reading comprehension, the evaluation employed multiple-choice questions, with each question carrying a value of one mark. The scoring criteria for writing and speaking were derived from Cambridge rubrics (Coombe, 2012). The rubrics were discussed and followed by two raters (see Appendix D), when there was disagreement, it was resolved through discussion.

4.2.4. **Self-Diagnostic Survey.** The self-diagnostic survey was adopted from Wu et al. (2021) and translated into learners' L1 (Arabic), as depicted in Appendix E. It was introduced to participants, after completing the EIT. The participants were asked to respond to all three sections, which encompass: perceptions of their performance during the EIT; perceptions of their listening and speaking abilities; and history of learning English. The first section asked participants to report their perception of their EIT performance, and their self-evaluation of the extent to which their comprehension and production of EIT sentences were influenced by vocabulary, grammar, accent, and pace of speech. The second section pertains to participants' perceptions of their listening and speaking abilities, evaluated through two distinct methods: first, employing a 5-point Likert scale for self-rating their overall ability, with ratings ranging from 5 (Excellent) to 1 (Needs Improvement). The second method is a self-evaluation of 11 Can-Do statements, utilizing a scale where 5 indicated "Not at All Difficult" and 1 indicated "Extremely Difficult." (see Appendix E). A calculation of reliability for section 1 showed high internal consistency: Cronbach's alpha was 0.81. Likewise, the second section indicated strong internal consistency: Cronbach's alpha was 0.87. The self-diagnostic survey's 5-point Likert scale ratings were consolidated, and participants' evaluations of their proficiency in the 11 Can-Do statements were tallied, resulting in a possible total score for speaking that fell within the range of 11 to 55 points.

4.3. Procedures

The OPT was introduced to all participants within a classroom setting (details of the OPT can be found in section 6.1). This was followed by the scheduling of individual meetings where the EIT was individually administered to each participant in a separate room (refer to section 6.1 and Appendix B). Following the completion of the EIT assessments, participants were presented with a self-diagnostic survey (refer to section 6.4 and Appendix E). In addition, the students' language scores were subsequently collected at the end of the semester. The dataset encompassed their scores in the final exams of language skills (GPA, speaking, listening, vocabulary, and grammar) in the EAP course. Information about scoring and related criteria can be found in section 6.3.

5. Results

The study employs a predictive and correlational research design. First, to address the first research question regarding the predictive validity of the EIT and OPT on the language skills of EFL

students, multiple regression analyses were run with GPA, listening, speaking, grammar, and vocabulary as dependent variables and EIT and OPT as independent variables. The prerequisites for regression analysis were satisfied in the current data. To check the multicollinearity assumption, previous research has proposed that VIF (variance inflation factor) exceeding 10 or tolerance values falling below 0.10 could signal the presence of multicollinearity and warrant further examination (Pallant, 2016). In the current study, the correlation's VIF is 1.95, and the tolerance is 0.51. These values indicate that the multicollinearity assumption has not been violated. The analysis encompasses various measures, including students' GPA and their scores in speaking, listening, vocabulary, and grammar. Descriptive statistics for both the independent and dependent variables are presented in Tables 1.

Table 1
Descriptive Statistics for Independent and Dependent Variables

	N	Minimum	Maximum	Mean	Standard Deviation
EIT	41	20	96	95.6	19.6
OPT	41	8	50	21	8.4
GPA	41	56.5	88.5	77.4	7.6
Speaking	41	3	15	10.7	3.1
Listening	41	10.5	15	13.02	1.03
Vocabulary	41	6	15	12.80	2.18
Grammar	41	6	15	12.44	2.10

Preliminary analysis indicated that there was a strong correlation between the two tests (r = .70). The shared variance between these two tests is about 49 % (R^2 = .70) which means that 49% of the variance in OPT scores can be explained by participants' EIT performance.

To answer the first research question, which examined whether EFL students' scores in GPA and language skills can be predicted by EIT and OPT, a multiple regression was conducted with GPA and language skills as the dependent variables and EIT and OPT scores as independent variables. Table 2 demonstrates that all the models predicting EFL students' scores from EIT and OPT scores achieved statistically significant levels.

Table 2
Multiple Regressions Models Predicting EFL Students' Achievement from OPT and EIT

Measures	Predictors	B	SE	β	*F*	*p*	R square	Adjusted R^2	Effect size Adjusted R^2 (Cohen's f^2)
GPA	Intercept	59.31	2.66			.000			
	EIT	.249	.058	.64	F (2,	.000	.57	.55	1.32
	OPT	.147	.136	.16	38) =	.287			
					25.3				
					2				
Speaking	Intercept	3.76	1.12		F (2,	.000			
	EIT	.081	.025	.51	38) =	.002	.53	.50	1.13
	OPT	.099	.057	.27	21.3	.093			
					2				
Listening	Intercept	12.29	.504		F (2,	.001			
	EIT	.044	.011	.76	38) =	.000	.30	.26	.42
	OPT	-.067	.026	-.498	8.10	.013			
Vocabulary	Intercept	8.50	.888		F (2,	.000			
	EIT	.058	.019	.52	38) =	.005	.41	.37	.69
	OPT	.040	.045	.16	12.9	.381			
					5				

Grammar	Intercept	9.35	.958		F (2,	.006			
	EIT	.047	.021	.44	38) =	.032	.23	.19	.30
	OPT	.013	.049	.049	5.80	.785			

For the model predicting GPA from EIT and OPT, EIT scores made a significant contribution, F (2,38) = 25.32, $p < .000$, whereas OPT scores did not. The finding indicates that this model explains 55% of the variance in GPA (adjusted R^2 = .55). The significant predictor in this model was the EIT, contributing 64%, while the OPT, although not statistically significant, contributed 16% to the model. The assessment of the significance of adjusted R^2 values followed the framework introduced by Plonsky and Ghanbar (2018). According to their categorization, values below .20 are classified as small, while those exceeding .50 are designated as large. This means that this model indicates strong predictability of the variance in EFL students' GPA, primarily driven by the EIT scores.

As for the other language skills, Table 6 indicates that for speaking, the model reached significance, F (2,38) = 21.32, $p < .000$, and predicted 50% of the variance in speaking scores (adjusted R^2 = .50). The significant predictor in this model was the EIT, contributing 51%. Despite OPT not being a significant predictor, its contribution was approximately 27%. This suggests that EFL students' speaking scores can be considerably predicted by their EIT performance. As for listening, the model reached statistical significance, F (2,38) = 8.10, $p < .001$. EIT contributed 76% to this model, while OPT contributed 50% of the variance. It should be noted that OPT showed a negative value which means that students who achieved low scores in OPT performed well in listening. This is contrary to the expectation and will be further explained in the next section. The models for vocabulary and grammar also reached statistical significance, F (2,38) = 12.95, $p < .000$; F (2,38) = 5.80, $p < .006$, respectively, with the EIT as the strong contributor in both models, accounting for .52 and .44 of the variances. These models account for 37% and 19% of the variance in EFL students' scores in vocabulary and grammar respectively. To sum up, the findings suggest that EIT considerably predicted EFL students' achievement in terms of GPA, speaking, listening, vocabulary, and grammar.

To answer the second research question, the Pearson product-moment correlation coefficient was conducted (Table 4). This research question investigates the extent to which students' perception of their listening and speaking ability aligns with their EIT performance. Two assessment methods were utilized to gauge students' perception of each ability: overall assessment and self-evaluation based on 11 Can-Do statements for both listening and speaking abilities (see Appendix E). The interpretation of significant correlations followed Cohen's (1988) guidelines, where r values between 0.10 and 0.29 were categorized as small, those between 0.30 and 0.49 were considered medium, and values between 0.50 and 1.0 were classified as large. Descriptive statistics can be found in Table 3.

Table 3
Descriptive Statistics of Perceived Listening and Speaking Abilities

Students' perception	N	Minimum	Maximum	Mean	SD
Perceived listening (overall ability)	41	1	5	3.23	1.34
Perceived listening (Can-do statements)	41	30	55	42.24	9.21
Perceived speaking (overall ability)	41	1	5	2.98	1.55
Perceived speaking (Can-do statements)	41	27	53	42.20	9.43

Table 4
Pearson's Correlations between Perceived Listening and Speaking Abilities and EIT Performance

		Perceived listening (overall)	Perceived listening (Can-do-statements)	Perceived speaking (overall)	Perceived speaking (Can-do-statements)
EIT	r	.479**	.447**	.530**	.509**
41	Sig.	.002	.004	.001	.001

**. Correlation is significant at 0.01; *. Correlation is significant at 0.05

Regarding perceived listening ability, Table 4 reveals a statistically significant positive correlation between the EIT scores and students' perception of their overall listening ability, $r = .479$, $p = .002$, as well as their self-evaluation based on Can-Do statements related to listening, $r =. 447$, $p = .004$. These correlations are considered moderate according to Cohen (1988) benchmark, indicating a moderate alignment between students' self-perceived listening ability and their performance on the EIT. Additionally, Table 4 demonstrates a significant positive correlation between EIT performance and perceived overall speaking ability, $r = .530$, $p = .001$, as well as with students' self-assessment based on Can-Do statements for speaking, $r = .509$, $p = .001$. These correlations are viewed as strong, suggesting a robust association between students' perceptions of their speaking ability and their performance on the EIT.

The third research question examined the extent to which students' EIT performance aligns with their perception of their EIT performance and the factors that affect their comprehension and repetition of sentences. Descriptive statistics for students' perceived performance and their evaluation of the factors impacting their comprehension and production of EIT sentences are presented in Table 5.

Table 5
Descriptive Statistics of Perceived Performance and Factors Affecting Comprehension and Production

Measures	Mean	Median	SD	Minimum	Maximum
Perceived performance	3.15	3	1.12	1	5
Factors affecting Comprehension					
Vocabulary	3.4	3.5	1.06	1	5
Length	3.08	3	1.16	1	5
Grammar	4.03	4	.89	2	5
Lack of context	4.37	5	.92	1	5
Pace of speech	3.23	3	1.21	1	5
Comprehension time	3.37	3.50	1.29	1	5
Accent	3.50	4	1.30	1	5
Factors affecting Comprehension					
Partial comprehension	4.98	5	.16	4	5
Length	4.93	5	.27	4	5
Grammar	4.75	5	.54	3	5
Pronunciation	4.70	5	.52	3	5
Retention of details	4.05	4	.90	2	5
Production time	3.75	4	1.15	1	5

The Pearson product-moment correlation coefficient was also conducted to answer this question. In terms of perceived performance, the correlation analysis revealed a positive correlation between EIT scores and students' self-perceived EIT performance, demonstrating a significant correlation of $r = .42$, $p = .007$. This implies that students' own assessments of their EIT performance closely mirror their actual EIT scores. Essentially, it suggests that students' self-perceptions regarding their performance in the EIT are somewhat accurate, as they align positively with their actual EIT scores. On the other hand, Table 6 presents a correlation analysis between EIT performance and students' self-assessment of factors influencing their comprehension of EIT sentences. The results suggest that there are moderate to strong correlations between EIT performance and most of these factors. Specifically, accent, sentence length, and vocabulary exhibited particularly strong correlations with EIT performance ($r = .597$, .550, and .551, respectively). These findings suggest that factors such as accent, sentence length, and vocabulary likely influence students' comprehension of EIT sentences and, in turn, impact their EIT performance. It is important to note that while these correlations are strong, they do not establish a causal relationship.

Table 6
Pearson's Correlations between EIT Performance and Factors Affecting Comprehension

		Vocabulary	Length	Grammar	Lack of context	Pace of speech	Comprehension time	Accent
EIT	r	.551**	.550**	.472**	.361*	.225	.458**	.597**
41	Sig.	.000	.000	.002	.022	.164	.003	.000

**. Correlation is significant at 0.01; *. Correlation is significant at 0.05

Table 7 indicates that EIT scores significantly correlated with specific factors that influence students' ability to repeat EIT items, particularly factors such as grammar, pronunciation, and retention of details. However, among these factors, only retention of details showed a strong correlation with EIT performance. This implies that the ability to remember specific details from the sentences has a significant impact on the production of these sentences. While other factors, such as grammar and pronunciation, may contribute, their impact is comparatively less significant.

Table 7
Pearson's Correlations between EIT Performance and Factors Affecting Production

		Partial comprehension	Length	Grammar	Pronunciation	Retention of details	Production time
EIT	r	.110	.152	.335*	.374*	.508**	.263
41	Sig.	.499	.349	.035	.017	.001	.101

**. Correlation is significant at 0.01; *. Correlation is significant at 0.05

In short, the correlation findings between students' scores on EIT and their perceived difficulty regarding factors influencing comprehension and production of EIT performance suggest that factors affecting comprehension may be of greater concern to EFL students. This is indicated by the significant correlations between most of these factors and their EIT performance, highlighting that EFL students tend to place more emphasis on comprehension-related factors.

6. Discussion

This study aims to evaluate the predictive validity of the EIT in an EFL classroom. The features of the EIT performance in the EFL classroom align with previously reported results in terms of reliability, with a coefficient of $\alpha = .95$, consistent with findings from previous studies (see Gaillard & Tremblay, 2016; Kim et al., 2016; Wu et al., 2021; Yan et al., 2016). Furthermore, the EIT demonstrates robust external validity, as indicated by its significant correlation with the OPT ($r = .70$), which assesses grammar and vocabulary in a written modality. This supports the assumption that EIT is modality independent (see Wu et al., 2021 for further discussion). Furthermore, the positive correlation between EIT scores and students' self-assessed EIT performance indicates that students' own perceptions align with their actual EIT scores ($r = .42$), further enhancing the test's validity. This section provides a discussion of the findings in relation to the research questions and the existing body of literature.

6.1. Predicting EFL Students' Achievement from EIT and OPT

The first research question aimed to investigate the predictive validity of EFL students' achievement based on their scores in the EIT and OPT. To answer this question, a multiple regression analysis was conducted. The results of the multiple regression analysis suggest that EIT scores play a significant role in predicting EFL students' GPA and other language skills, while OPT scores have a limited predictive value in this context. The findings indicated that the EIT demonstrated substantial predictive validity for various aspects of EFL students' achievement, including GPA, speaking, listening, vocabulary, and grammar. It appears to be a robust tool for predicting these language skills in the context of the study. However, it is important to note the unexpected negative relationship between OPT scores and listening scores. It is possible that some students may have experienced anxiety at the

beginning of the semester when the OPT was administered which may had an impact on their performance on the OPT. EIT, on the other hand, may not trigger the same level of anxiety for EFL students as they may feel less intimidated when asked to repeat as much of each sentence they hear (see Wu & Ortega, 2013). However, this assumption needs to be further explored in future research.

The results of this study support earlier research indicating that the EIT not only measures communication skills in speaking and listening but also has the capacity to assess other aspects of language proficiency, such as grammar and vocabulary, (see Gass, 2018; Wu & Ortega, 2013). It has been suggested that EIT performances provide valuable insights into learners' command of vocabulary and grammar, as well as their ability to use them fluently and with reasonably clear pronunciation (Wu & Ortega, 2013). This finding suggests that EIT scores can effectively predict students' overall English proficiency, encompassing skills beyond oral production, such as English GPA which includes reading and writing. Future research within EFL classrooms is needed to validate the current findings.

6.2. The Relationship between Students' Perceived Listening and Speaking Abilities and Their EIT Performance

The second research question explored the relationship between students' self-perceptions of their listening and speaking abilities and their EIT performance. Two methods were employed to assess each skill, including students' perceptions of their overall ability and their evaluation of each skill based on 11 Can-Do statements (see Appendix E). The results reveal that students' perceived listening ability, evaluated through two methods, demonstrates a moderate statistically significant correlation with EIT scores. Similarly, their perceived speaking ability indicates a strong correlation with EIT scores when assessed through two different methods. This finding carries significant implications, suggesting that students' self-perception of their listening and speaking abilities reasonably align with their EIT performance. The significance of this result extends beyond the mere correlation between perceived listening and speaking abilities and EIT performance. It underscores the relevance of students' self-perceptions in understanding their language proficiency, particularly in the EIT. It indicates that students' self-awareness of their speaking skills, in particular, can closely mirror their ability to comprehend and imitate spoken language.

The current findings align with previous research conducted within EFL contexts. For example, Bowden (2016) found a significant correlation between perceived listening and EIT scores, as well as between perceived speaking skills and EIT performance. In another EFL context, Wu et al. (2021) found that participants' self-assessment of their listening skills, rather than speaking skills, significantly predicted EIT performance across two parallel EIT forms. It has been argued that in many EFL environments, learners often demonstrate better receptive listening skills compared to their productive speaking skills, given that integrating listening training in foreign language classrooms is typically more feasible than prioritizing speaking skills (Wu et al., 2021). However, this assumption may not hold true in the context of the present study, as EFL learners displayed a notable awareness of both their listening and speaking abilities. One potential explanation for the current findings could be that in the study's context, EFL students were enrolled in an intensive EAP course, where equal emphasis was placed on developing listening and speaking skills. Consequently, this balanced approach may have facilitated the development of a more accurate perception of both listening and speaking abilities among the students.

In conclusion, the results obtained from the investigation of the second research question offer valuable insights into the relationship between students' self-perceptions and their EIT performance. The significance of the perceived listening and speaking skills underscores the multifaceted aspect of language assessment, wherein learners' self-perceived abilities can serve as an assessment resource.

6.3. Perception of the Factors Affecting Comprehension and Production

The third research question examined the extent to which students' EIT performance aligns with their perception of the factors that affect their comprehension and repetition of sentences. The results highlight the significant association between students' EIT scores and comprehension factors such as accent, sentence length, and vocabulary. The findings suggest that these factors might be key determinants of how well students grasp the sentences presented in the EIT. In contrast, factors like grammar, comprehension time, and the absence of context have a relatively moderate association with EIT scores. Concerning accents, the present study involved EFL learners who, in their daily lives, were

not exposed to a wide range of English varieties. The participants commented that they encountered difficulties when listening to the British accent used in the EIT. This observation aligns with the research conducted by Wu et al. (2021), which emphasized that the speaker's pronunciation plays a significant role in determining the level of difficulty experienced by EFL learners within their study's specific context. This implies that within the EFL context, the speaker's pronunciation style, including accent-related features, can substantially impact learners' ability to comprehend and reproduce spoken language.

The other factors which also indicated a strong correlation with EIT performance in this study include sentence length and vocabulary. It is widely recognized that prompt length is a strong predictor of variances in EIT scores (see Vinther, 2002; Wu & Ortega, 2013; Yan et al., 2016). This suggests that students' perceptions of sentence length as a factor of difficulty in comprehending EIT prompt align with previous research findings. Likewise, vocabulary has been found to explain part of the score variance in previous studies (e.g., Graham et al. 2010; Tracy-Ventura et al., 2014). In fact, Graham et al. (2010) highlighted the importance of considering lexical difficulty when designing sentences for EITs. It has been suggested that the influence of lexical frequency is particularly notable when sentences approach the limits of working memory (Graham et al., 2010). It appears that a strong command of vocabulary may be crucial for understanding EIT prompts, similar to what was observed in previous literature (Graham et al., 2010; Wu & Ortega, 2013). Other factors, such as grammar, comprehension time, and the absence of context, which were moderately correlated with students' perceptions of difficulty, have also been discussed in the literature as contributing to increased difficulty in the EIT (see Gass, 2018; Tracy-Ventura et al., 2014; Vinther, 2002; Wu & Ortega, 2013; Yan et al., 2016).

Concerning production aspects, only the retention of specific details displayed a strong correlation with EIT performance. This implies that the ability to recall specific details from the sentences may have a significant association with the accurate repetition of these sentences according to students' evaluation. This means that participants who rated the retention of specific details as a strong obstacle to production also achieved higher scores on the EIT, indicating that their memory might not fully support them in repeating specific sentence details, but their high proficiency level compensates for this. This implies that language proficiency appears to play a considerable role in performing well on EIT despite potential memory limitations.

The findings for this research question suggest that factors relating to comprehension might be more prevalent in the minds of EFL learners than factors relating to the production of EIT. This is similar to the findings of Wu et al. (2021), as comprehension stood out prominently as a major concern for beginners and intermediate proficiency learners, and it contributed significantly to the EIT scores in the multiple regression analysis.

7. Conclusion

The present study underscores the distinctiveness of EIT as an assessment tool that goes beyond merely evaluating language proficiency; it can also serve as a valuable predictor of students' overall linguistic skills in the EFL context. In practical terms, these findings signify that educators and institutions can potentially use EIT scores as an effective tool for identifying students who may be facing academic challenges in their EFL courses. Additionally, the current findings indicate that learners' self-evaluations can be effectively utilized as an assessment resource. The findings also imply that comprehension-related factors may be more prominent and challenging for EFL learners compared to those related to EIT production.

The present study, nonetheless, has some limitations that need to be acknowledged. First, this study was conducted within a specific EFL classroom context, which might limit the generalizability of the findings to other educational or research settings. Replicating the study in various EFL contexts could provide a more comprehensive understanding of the factors at play. In addition, the current study has a small sample size which might not adequately represent the broader population or context that the study aims to investigate. Therefore, caution should be exercised when attempting to apply the study's results to larger or more diverse populations.

Notwithstanding these limitations, the study provided novel insights into the predictive validity of the EIT within EFL classrooms. It also underscored the link between students' self-perceived speaking and listening abilities and their actual EIT performance. Furthermore, the present study

illuminated the degree to which factors associated with comprehension and production relate to students' performance in the EIT. However, there are specific areas that warrant attention in future research. First, future research should delve into the potential impact of students' proficiency levels on their perceptions of factors influencing comprehension and production in the EIT. Understanding how students of varying proficiency levels perceive and navigate the challenges related to comprehension and production in the EIT is crucial. Finally, future research needs to investigate specific individual factors associated with test-taking, particularly focusing on the levels of anxiety experienced during EIT in comparison to other assessments employing different modalities.

Declaration of Conflicting Interests
The author declared no potential conflicts of interest regarding the research, authorship, and/or publication of this article.

Funding
The author received no financial support for the research, authorship, and/or publication of this article.

References

Akbary, M., Benzaia, L. A., Jarvis, S., & Park, H. I. (2023). Evaluating the utility of elicited imitation as a measure of listening comprehension in the context of forensic linguistics. *Research Methods in Applied Linguistics*, *2*(3), 100067. https://doi.org/10.1016/j.rmal.2023.100067

Albarqi, G. (2024). Predicting elicited imitation performance from complexity, accuracy and fluency (CAF) of L2 oral production. *International Journal of Language and Literary Studies*, *6*(1), 226–242. https://doi.org/10.36892/ijlls.v6i1.1569

Allan, D. (2004). *Oxford placement test. University of Cambridge local examination syndicate*. Oxford University Press. https://shorturl.at/qvIU6

Ariamanesh, A. A., Barati, H., & Youhanaee, M. (2023). TOEFL iBT speaking subtest: The efficacy of preparation time on test-takers' performance. *International Journal of Language Testing*, *13*(2), 38-55. https://doi.org/10.22034/IJLT.2022.357001.1189

Bacon, D. R. & Bean. B. (2006). GPA in research studies: An invaluable but neglected opportunity. *Journal of Marketing Education*, *28*(1). 35–42. https://doi.org/10.1177/0273475305284638

Bley-Vroman, R., & Chaudron, C. (1994). Elicited imitation as a measure of second-language competence. In E. E. Tarone, S. M. Gass, & A. D. Cohen (Eds.), *Research methodology in second-language acquisition* (pp. 245–261). Lawrence Erlbaum.

Bowden, H. W. (2016). Assessing second-language oral proficiency for research: The Spanish elicited imitation task. *Studies in Second Language Acquisition*, *38*, 647–675. https://doi.org/10.1017/S0272263115000443

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum.

Coombe, C. (Ed.). (2012). *The Cambridge guide to second language assessment*. Cambridge University Press.

Daller, M. H. & Phelan. D. (2013). Predicting international student study success. *Applied Linguistics Review*, *4*(1). 173–193. https://doi.org/10.1515/applirev-2013-0008

Daller, M., & Yixin, W. (2017). Predicting study success of international students. *Applied Linguistics Review*, *8*(4), 355-374. https://doi.org/10.1515/applirev-2016-2013

Davis, L., & Norris, J. (2021). Developing an innovative elicited imitation task for efficient English proficiency assessment. *ETS Research Report Series*, *2021*(1), 1–30. https://doi.org/10.1002/ets2.12338

Dörnyei, Z., & Chan, L. (2013). Motivation and vision: An analysis of future L2 self-images, sensory styles, and imagery capacity across two target languages. *Language Learning*, *63*(3), 437–462. https://doi.org/10.1111/lang.12005

Dörnyei, Z., & Katona, L. (1992). Validation of the C-test amongst Hungarian EFL learners. *Language Testing*, *9*(2), 187-206. https://doi.org/10.1177/026553229200900206

Duran-Karaoz, Z., & Tavakoli, P. (2020). Predicting L2 fluency from L1 fluency behavior: The case of L1 Turkish and L2 English speakers. *Studies in Second Language Acquisition*, *42*(4), 671-695. https://doi.org/10.1017/S0272263119000755

Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition*, *27*, 141–172. doi:10.1017/S0272263105050096

Erlam, R. (2006). Elicited imitation as a measure of L2 implicit knowledge: An empirical validation study. *Applied Linguistics*, *27*(3), 464–91. https://doi.org/10.1093/applin/aml001

Feast, V. (2002). The impact of IELTS scores on performance at university. *International Education Journal*, *3*(4). 70–85. https://shorturl.at/glxOQ

Gaillard, S., & Tremblay, A. (2016). Linguistic proficiency assessment in second language acquisition research: The elicited imitation task. *Language Learning*, *66*, 419–447. https://doi.org/10.1111/lang.12157

Gass, S. (2018). SLA elicitation tasks. In A. Phakiti, P. De Costa, L. Plonsky, & S. Starfield (Eds.), *The Palgrave Handbook of Applied Linguistics Research Methodology* (pp. 313–337). London: Palgrave Macmillan.

Gómez-Benito, J., Sireci, S., Padilla, J-L., Hidalgo, D., & Benítez, I. (2018). Differential item functioning: Beyond validity evidence based on internal structure. *Psicothema, 30*, 104–109. 9. https://doi.org/10.7334/psicothema2017.183

Graham, C. R., McGhee, J., & Millard, B. (2010). The role of lexical choice in elicited imitation item difficulty. In M. T. Prior, Y. Watanabe, & S.- K. Lee (Eds.), *Selected proceedings of the 2008 Second Language Research Forum: Exploring SLA perspectives, positions, and practices,* (pp. 57–72). Somerville, MA: Cascadilla Proceedings Project. https://www.lingref.com/cpp/libraries/index.html

Hood, L., & Lightbown, P. (1978). What children do when asked to "say what I say": Does elicited imitation measure linguistic knowledge? *Allied Health and Behavioral Sciences*, *1*, 195–220. https://doi.org/10.1177/0265532215594643

Isbell, D. R., & Son, Y. A. (2022). Measurement properties of a standardized Elicited Imitation Test: An integrative data analysis. *Studies in Second Language Acquisition*, *44*(3), 859-885. https://doi.org/10.1017/S0272263121000383

Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measuremen*t, *50*, 1–73. https://doi.org/10.1111/jedm.12000

Kim, Y. K., & Kim, T. Y. (2011). The effect of Korean secondary school students' perceptual learning styles and ideal L2 self on motivated L2 behavior and English proficiency. *Korean Journal of English Language and Linguistics*, *11*(1), 21–42. https://doi.org/10.15738/kjell.11.1.201103.21

Kim, H., Song, C., Kim, J., Jeong, H., & Park, J. (2024). Validation of an elicited Imitation test as a measure of Korean language proficiency. *Language Testing in Asia*, *14*(1), 17. https://doi.org/10.1186/s40468-024-00287-z

Kim, Y., Tracy-Ventura, N., & Jung, Y. (2016). A measure of proficiency or short-term memory? Validation of an elicited imitation test for SLA research. *Modern Language Journal*, *100*, 655–673. https://doi.org/10.1111/modl.12346

Kostromitina, M., & Plonsky, L. (2021). Elicited imitation tasks as a measure of L2 proficiency: A meta-analysis. *Studies in Second Language Acquisition*, *44*(3), 886-911. https://doi.org/10.1017/S0272263121000395

Ortega, L., Iwashita, N., Norris, J. M., & Rabie, S. (2002, October). *An investigation of elicited imitation tasks in crosslinguistic SLA research*. Paper presented at the Second Language Research Forum, Toronto, Ontario, Canada, October 3–6, 2002 [Unpublished handout retrieved from IRIS].

Park, H. I., Solon, M., Henderson, C., & Dehghan-Chaleshtori, M. (2020). The roles of working memory and oral language abilities in elicited imitation performance. *Modern Language Journal*, *104*, 133–151. https://doi.org/10.1111/modl.12618

Pallant, J. (2016). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS*. Berkshire: McGraw-Hill Education.

Plonsky, L., & Ghanbar, H. (2018). Multiple regression in L2 research: A methodological synthesis and guide to interpreting R2 values. *The Modern Language Journal*, *102*(4), 13-731. https://doi.org/10.1111/modl.12509

Révész, A., & Brunfaut, T. (2021). Validity in language assessment. In P. Winke & T. Brunfaut (Eds.) *The Routledge handbook of SLA and language assessment* (pp. 21–32). Routledge.

Skehan, P., Foster, P., & Shum, S. (2016). Ladders and snakes in second language fluency. *International Review of Applied Linguistics in Language Teaching*, *54*(2), 97-111. https://doi.org/10.1515/iral-2016-9992

Solon, M., & Park, H. I. (2024). Elicited imitation in second language acquisition research: New insights to advance methodological rigor (Introduction to the special issue). *Research Methods in Applied Linguistics*, *3*(2), 100112. https://doi.org/10.3138/cmlr.64.1.215

Souzandehfar, M. (2024). New perspectives on IELTS authenticity: An evaluation of the speaking module. *International Journal of Language Testing*, *14*(1), 34-55. https://doi.org/10.22034/IJLT.2023.409599.1272

Suzuki, Y., & DeKeyser, R. (2015). Comparing elicited imitation and word monitoring as measures of implicit knowledge. *Language learning*, *65*(4), 860-895. https://doi.org/10.1111/lang.12138

Tracy-Ventura, N., McManus, K., Norris, J., & Ortega, L. (2014). Repeat as much as you can: Elicited imitation as a measure of oral proficiency in L2 French. In P. Leclercq, A. Edmonds, & H. Hilton (Eds.), *Measuring L2 proficiency: Perspectives from SLA* (pp. 143–166). Multilingual Matters. https://doi.org/10.21832/9781783092291-011

Tremblay, A. (2011). Proficiency assessment standards in second language acquisition research: "Clozing" the gap. *Studies in Second Language Acquisition*, *33*, 339–372. https://doi.org/10.1017/S0272263111000015

Van Moere, A. (2012). A psycholinguistic approach to oral language assessment. *Language Testing*, *29*(3), 325-344. https://doi.org/10.1177/0265532211424478

Vinther, T. (2002). Elicited imitation: A brief review. *International Journal of Applied Linguistics*, *12*, 54–73. https://doi.org/10.1111/1473-4192.00024

Wongtrirat, Rachawan. *English Language Proficiency and Academic Achievement of International Students: A Meta-Analysis* (2010). Doctor of Philosophy (PhD), Dissertation, Educational Foundations & Leadership, Old Dominion University, DOI: 10.25777/y7yt-m587. https://digitalcommons.odu.edu/efl_etds/180

Woodrow, L. (2006). The academic success of international postgraduate education students and the role of English proficiency. *University of Sydney Papers in TESOL*, *1*(1), 51–70. https://shorturl.at/itMO6

Wu, S.-L., & Ortega, L. (2013). Measuring global oral proficiency in SLA research: A new elicited imitation test of L2 Chinese. *Foreign Language Annals*, *46*, 680–704. https://doi.org/10.1111/flan.12063

Wu, S. L., Tio, Y. P., & Ortega, L. (2021). Elicited imitation as a measure of L2 proficiency: New insights from a comparison of two L2 English parallel forms. *Studies in Second Language Acquisition*, *44*(1), 271-300. https://doi.org/10.1017/S0272263121000103

Yan, X., Maeda, Y., Lv, J., & Ginther, A. (2016). Elicited imitation as a measure of second language proficiency: A narrative review and meta-analysis. *Language Testing*, *33*, 497–528. https://doi.org/10.1177/0265532215594643

Yu, Q. (2014). *Various items causing IELTS test-taker's low performance in mainland China: An international joint education program solution*. Paper presented at the 2014 International Conference on Global Economy, Finance and Humanities Research (GEFHR 2014). Paris, France: Atlantis Press. https://doi.org/10.2991/gefhr-14.2014.7

Zumbo, B. D. (2007). Three generations of DIF analyses: Considering where it has been, where it is now, and where it is going. *Language Assessment Quarterly*, *4*, 223–233. https://doi.org/10.1080/15434300701375832

Appendix A

Proficiency Levels According to the OPT scores

CEFR levels	OPT scores	Number of participants
C2	55 -60	0
C1	48-54	0
B2	40-47	0
B1	30-39	5
A2	18-29	19
A1	0-17	15

Appendix B

EIT Sentences (Wu & Ortega, 2013)

1. I have to get a haircut. (7)
2. The red book is on the table. (8)
3. The streets in this city are wide. (8)
4. He takes a shower every morning. (9)
5. What did you say you were doing today? (10)
6. I doubt that he knows how to drive that well. (10)
7. After dinner I had a ling, peaceful nap. (11)
8. It is possible that it will rain tomorrow. (12)
9. I enjoy movies which have a happy ending. (12)
10. The houses are very nice but too expensive. (12)
11. The little boy whose kitten died yesterday is sad. (13)
12. That restaurant is supposed to have very good food. (13)
13. I want a nice, big house in which my animals can live. (14)
14. You really enjoy listening to country music, don't you? (14)
15. She just finished painting the inside of her apartment. (14)
16. Cross the street at the light and then just continue straight ahead. (15)
17. The person I'm dating has a wonderful sense of humor. (15)
18. She only orders meat dishes and never eats vegetables. (15/16)
19. I wish the price of town houses would become affordable. (15)
20. I hope it will get warmer sooner this year than it did last year. (16)
21. A good friend of mine always takes care of my neighbor's three children. (16)
22. The black cat that you fed yesterday was the one chased by the dog. (16)
23. Before he can go outside, he has to finish cleaning his room. (16)
24. The most fun I've ever had was when we went to the opera. (16)
25. The terrible thief whom the police caught was very tall and thin. (17)
26. Would you be so kind as to hand me the book which is on the table? (17)
27. The number of people who smoke cigars is increasing every year. (17/18)
28. I don't know if the 11:30 train has left the station yet. (18)
29. The exam wasn't nearly as difficult as you told me it would be. (18)
30. There are a lot of people who don't eat anything at all in the morning. (19)

Appendix C

EIT Scoring Rubric (based on Ortega et al., 2002)

Item Score	Description	Examples
4	Perfect repetition	-That restaurant is supposed to have very good food -That restaurant is supposed to have very good food
3	Accurate content repetition with some changes of form	-The houses are very nice but too expensive (12) -The houses are very nice but it expensive
2	Changes in content or in form that affect meaning	-It is possible that it will rain tomorrow -It is impossible to train tomorrow
1	Repetition of half or less of the stimulus leading to substantial loss of meaning	- The little boy whose kitten died yesterday is sad. - The little boy whose kitten
0	Silence, only one word repeated, or unintelligible repetition	-No response -The boy

Note. Examples are taken from the data in the current study.

Appendix D

The Scoring Criteria for Writing and Speaking based on Cambridge Assessment (Coombe, 2012)

Cambridge Speaking Assessment Scales

Quick guide to marking

A mark of **5** means that the learner's speaking is well above Pre-intermediate level.

A mark of **3** means that the learner's speaking is average for Pre-intermediate level.

A mark of **1** means that the learner's speaking is not coherent enough to be properly assessed.

	Task achievement	Range	Organisation	Pronunciation	Accuracy
	Did the learner complete the tasks in each section of the test?	*Did the learner use a wide variety of vocabulary and grammar structures?*	*Did the learner express his/her ideas clearly and connect them together effectively?*	*Did the learner produce individual sounds clearly and use appropriate stress and intonation?*	*Did the learner produce grammatically correct language?*
5	All tasks completed successfully and without difficulty	A wide variety of both appropriate structures and vocabulary used confidently	Learner able to connect ideas clearly and effectively, using basic linkers and devices appropriate to the level	Clear pronunciation of sounds and use of stress and intonation to convey meaning, though L1 accent may be discernible	Very few errors in use of structures and vocabulary expected at this level, perhaps with evident ability to self-correct
4	Most tasks completed without difficulty	Some elements of 3 and some of 5	Some elements of 3 and some of 5	Some elements of 3 and some of 5	Some elements of 3 and some of 5
3	Tasks completed with occasional difficulty	A variety of appropriate structures used, with some inappropriate usage or difficulty retrieving appropriate language	Learner usually able to communicate and link ideas clearly, though with a lack of overall fluency	Some difficulty with sounds, stress and intonation, causing strain on the listener, though communication is rarely impaired	Errors in use of structures and vocabulary are frequent, but rarely impair communication
2	Frequent difficulties prevented task completion	Some elements of 1 and some of 3	Some elements of 1 and some of 3	Some elements of 1 and some of 3	Some elements of 1 and some of 3
1	Too little communication to assess	Too little communication to assess	Too little communication to assess	Too little communication to assess	Too little communication to assess

Cambridge Writing Assessment Scales

Quick guide to marking

A mark of **5** means that the learner's writing is well above Starter level.

A mark of **3** means that the learner's writing is average for Starter level.

A mark of **1** means that the learner's writing is not coherent enough to be properly assessed.

	Task achievement	Range	Organisation	Register	Accuracy
	Did the learner complete the tasks in each section of the test?	*Did the learner use a wide variety of vocabulary and grammar structures?*	*Did the learner express his/her ideas clearly and connect them together effectively?*	*Did the learner use language with an appropriate style and level of formality?*	*Did the learner produce grammatically correct language?*
5	All tasks completed successfully and without difficulty	A wide variety of both correct structures and appropriate vocabulary used	Learner able to connect ideas clearly and effectively, using basic linkers and devices appropriate to the level	The learner is able to use appropriate language when writing for different reasons and at different levels of formality	Very few errors in use of structures and vocabulary expected at this level
4	Most tasks completed without difficulty	Some elements of 3 and some of 5	Some elements of 3 and some of 5	Some elements of 3 and some of 5	Some elements of 3 and some of 5
3	Tasks completed with occasional difficulty	A variety of appropriate structures used, with some inappropriate or repetitive use of language	Learner usually able to communicate and link ideas clearly, though sometimes errors make meaning unclear	Meaning is largely clear, but the learner is not able to choose language according to register	Errors in use of structures and vocabulary are frequent, but rarely impair communication
2	Frequent difficulties prevented task completion	Some elements of 1 and some of 3	Some elements of 1 and some of 3	Some elements of 1 and some of 3	Some elements of 1 and some of 3
1	Too little communication to assess	Too little communication to assess	Too little communication to assess	Too little communication to assess	Too little communication to assess

Appendix E

The Self-diagnostic Survey based on Wu et al., 2021 (translated into Arabic)

Name: _____ الاسم:

Consent: I give consent to participate in this survey (Yes/ No). الموافقة: أوافق على المشاركة (نعم/ لا)

Regarding the English Repetition Task فيما يتعلق باختبار تكرار الجمل

1. How do you evaluate the overall level of difficulty of the repetition task?
   كيف وجدتي مستوى صعوبة الجمل
   5: Extremely difficult صعب جدا
   4: Quite difficult صعب
   3: Somewhat difficult صعب نوعا ما
   2: Slightly difficult صعب قليلا
   1: Not at all difficult ليس صعب اطلاقا
2. How do you evaluate your overall performance on the repetition task?
   كيف تقيمين آدائك في الاختبار
   5: Excellent ممتاز
   4: Good جيد
   3: Average متوسط
   2: Fair مقبول
   1: Needs improvement بحاجة الى تحسين
3. Please use the following scale to rate how frequently each statement below affects your performance on the repetition task.
   رجاء استخدمي المقياس التالي لتقدير مدى تأثير كل جملة على ادائك اثناء تكرار الجمل
   5: almost always دائما
   4: frequently في أغلب الاوقات
   3: sometimes بعض الاوقات
   2: occasionally قليلا
   1: never اطلاقا

   ____ I find some sentences hard to understand because there are vocabulary words I don't know.
   وجدت صعوبة في فهم بعض الجمل لان هناك مفردات لاأعرفها

   ____ I find some sentences hard to understand because the sentences are too long to follow.
   وجدت صعوبة في فهم بعض الجمل لان هناك جمل طويلة جدا.

   ____ I find some sentences hard to understand because the grammar/structure throws me off.
   وجدت صعوبة في فهم بعض الجمل لصعوبة القواعد

   ____ I find some sentences hard to understand because they are isolated sentences without context.
   وجدت صعوبة في فهم بعض الجمل لان الجمل وردت منفصلة وبدون سياق

   ____ I find some sentences hard to understand because the pace of speech is too fast for me to follow.
   وجدت صعوبة في فهم بعض الجمل لان النطق كان سريع جدا بالنسبة لي

   ____ I find some sentences hard to understand because of insufficient time for me to process the meaning.
   وجدت صعوبة في فهم بعض الجمل لان الوقت لم يكن كافي لاستيعاب المعنى

   ____ I find some sentences hard to understand because the accent and pronunciation are unfamiliar to me.
   وجدت صعوبة في فهم بعض الجمل لان اللكنة والنطق غير مألوف بالنسبة لي

   Other. Please elaborate_____
   اذا كان لديك إجابة أخرى اذكريها هنا
4. Please use the following scale to rate how frequently each statement affects you.
   رجاء استخدمي المقياس التالي لتقدير مدى تأثير كل جملة على ادائك اثناء تكرار الجمل
   5: almost always دائما
   4: frequently في أغلب الاوقات
   3: sometimes بعض الاوقات
   2: occasionally قليلا
   1: never اطلاقا

____ I find it hard to repeat because I don't understand a portion of some sentences I heard.
وجدت صعوبة في الإعادة لأنني لم أفهم جزء من الجمل

____ I find it hard to repeat because some sentences are too long to remember.
وجدت صعوبة في الإعادة لأن بعض الجمل طويلة ولم أتمكن من تذكرها

____ I find it hard to repeat because I have a hard time reproducing some grammar/structure.
لأنني واجهت وقتا عصيبا في إعادة بعض القواعد وجدت صعوبة في الإعادة

____ I find it hard to repeat because I have a hard time pronouncing some vocabulary words.
لأنني واجهت وقتا عصيبا في نطق بعض المفردات وجدت صعوبة في الإعادة

____ I find it hard to repeat because I remember the main idea of the sentence but forget the exact wording or the details of it.
لأنني أتذكر الفكرة الرئيسية للجملة لكنني نسيت الكلمات بالتحديد او تفاصيلها وجدت صعوبة في الإعادة

____ I find it hard to repeat because I speak too slow to repeat the sentence within the given time.
وجدت صعوبة في الإعادة لأنني اتحدث ببطء لأعيد الجملة في الوقت المحدد

Other. Please elaborate _____
اذا كان لديك إجابه أخرى ارجو كتابتها هنا

5. In this repetition task, which is more challenging to you, understanding the sentence or repeating the sentence? Why?
في تمرين التكرار،أيهما اصعب فهم الجمله او تكراراها، ولماذا؟
_____

Regarding Your English Language Ability فيما يتعلق بمستواك في اللغة الإنجليزية

6. Based on your own assessment, how would you rate your overall listening skills?
بناء على تقييمك الشخصي لمستواك اللغوي، كيف تقيمين مهارة الاستماع لديك؟

5: Excellent ممتاز
4: Good جيد
3: Average متوسط
2: Fair مقبول
1: Needs improvement بحاجة لتحسين

7. Please use the following scale to rate how well you can perform the following tasks in English.
ارجو استخدام المقياس التالي لتقييم مقدرتك اللغوية على آداء المهام التالية باللغة الإنجليزية:

5: not at all difficult ليست صعبة على الاطلاق
4: slightly difficult صعبة قليلا
3: somewhat difficult صعبة الى حد ما
2: quite difficult صعبة جدا
1: extremely difficult في غاية الصعوبة

___ I can understand greetings. استطيع فهم التحية
___ I can understand days of the week and the hour.استطيع فهم أيام الأسبوع والوقت
___ I can sometimes understand a simple transaction between a customer and a sales clerk.
استطيع أحيانا فهم محادثة بسيطة بين بائع ومشتري
___ I can understand questions about my likes and dislikes.
استطيع فهم الأسئلة التي تدور حول الأشياء التي أحبها او لا أحبها
___ I can understand that an event is being postponed or cancelled.استطيع فهم الجمل التي تتحدث عن تأجيل أو إلغاء أنشطة معينة.
___ I can understand a voice message from the airlines about changes to a flight schedule. أستطيع فهم رسالة صوتية من خطوط الطيران عن تعديل موعد الرحلة.
___ I can understand short presentations about famous people in history.
استطيع فهم عروض قصيرة تتحدث عن أشخاص لهم شهرة تاريخية.
___ I can understand an interview with a famous person, such as a rock star, politician, or actor.استطيع فهم مقابلة مع شخصية مشهورة مثل مغني، او سياسي او ممثل
___ I can understand a group leader's justification for protesting a cut in programs. استطيع فهم التبرير المقدم من قبل قائد احدى المجموعات حول الاعتراض على تقليص البرامج.
___ I can understand a speech on a historical period. .استطيع فهم خطاب يدور حول حقبة تاريخية معينة
___ I can understand the plot and cultural implications of oral narratives such as folk and fairy tales. استطيع فهم القصص والمعاني الثقافية لحكايات شفوية كالقصص الشعبية والأساطير

8. Based on your own assessment, how would you rate your overall speaking skills?
بناء على تقييمك الشخصي لمستواك اللغوي، كيف تقيمين مهارة التحدث لديك؟
5: Excellent ممتاز
4: Good جيد
3: Average متوسط
2: Fair مقبول
1: Needs improvement بحاجة لتحسين
9. Please use the following scale to rate how well you can perform the following tasks.
ارجو استخدام المقياس التالي لتقييم مقدرتك اللغوية على آداء المهام التالية باللغة الإنجليزية:
5: not at all difficult ليست صعبة على الاطلاق
4: slightly difficult صعبة قليلا
3: somewhat difficult صعبة الى حد ما
2: quite difficult صعبة جدا
1: extremely difficult في غاية الصعوبة
___ I can say hello and goodbye. استطيع ان ألقي التحية
___ I can introduce myself and provide basic personal information.
استطيع تقديم نفسي وإعطاء معلوماتي الشخصية
___ I can ask for directions to a place. استطيع ان أسأل عن اتجاه مكان معين
___ I can make a reservation. استطيع إجراء حجوزات
___ I can give reasons for my preferences. استطيع إبداء الأسباب لاختياراتي
___ I can talk about my family history. استطيع التحدث عن تاريخ عائلتي
___ I can explain how life has changed since I was a child and respond to questions on the topic.
استطيع الحديث عن التغيرات في حياتي منذ ان كنت صغيرة واستطيع الرد على الأسئلة التي توجه لي حول هذا الموضوع
___ I can discuss future plans, such as where I want to live and what I will be doing in the next few years. استطيع مناقشة خططي المستقبلية مثل أين اريد ان أعيش وماذا اريد ان افعل في المستقبل القريب
___ I can usually defend my views in a debate. استطيع أحيانا الدفاع عن آرائي اثناء النقاش
___ I can put forth and react to others' complex ideas during a business discussion.
استطيع ان استخدم و أقوم بالرد كذلك على حديث الآخرين والمشتمل على أفكار معقدة اثناء المناقشات الاقتصادية.
___ I can use my language persuasively to advocate a point of view that is not necessarily my own. استطيع استخدام لغة مقنعة لإيصال فكرة معينة

Regarding Your History of Learning the English Language فيما يتعلق بتاريخك في تعلم اللغة الإنجليزية
1. What is your major in college? _____ ماهو تخصصك؟
2. At what age did you start to learn English? _____ (years old)
_____ متى بدأتي تعلم اللغة الإنجليزية؟ كم عمرك وقتها؟
3. How long (in years) in total have you studied English at school? _____ (years)
منذ متى بدأتي تعلم اللغة الإنجليزية؟ ..... عاما
4. Have you visited/lived in an English-speaking country? □ No □ Yes (if YES, see below)
هل سبق وأن زرتي او عشتي في بلد يتحدث أهله الإنجليزية؟ نعم/لا (اذا اجبتي بنعم اكملي التالي)
I have been to _____ (name of the country) for _____ (week/month/year)
لقد عشت في _____ لمدة _____
I have been to _____ (name of the country) for _____ (week/month/year)
5. Use the following scale to rate how often you hear or use English in your daily life?
استخدمي المقياس التالي لتقييم مدة استماعك او استخدامك للغة الإنجليزية يوميا؟
5: almost always دائما
4: frequently غالبا
3: sometimes احيانا
2: occasionally قليلا
1: never اطلاقا