Attaining C1 Level Scores in TOEFL iBT: From Challenges to Needs

Omid S. Kalantar[1]*

ARTICLE INFO	ABSTRACT
	This study sought to identify the challenges and needs of TOEFL iBT candidates in achieving C1 level scores in the speaking and writing sections of the exam. To this end, the researcher employed a mixed-method approach to collect data from a population of 46 students, both male and female, between the ages of 22 and 30. The participants were enrolled in two TOEFL preparation classes at Universidad Nacional de Educación. The data were collected through a survey, class observations, and TOEFL speaking and writing mock tests, and were subsequently analyzed both qualitatively and quantitatively. The findings indicated that the participants encountered challenges in discussing unfamiliar topics fluently, struggled with producing language at the C1 level while employing a variety of grammar and vocabulary, and demonstrated reduced accuracy due to influences from their first language (L1). To address these challenges, it is recommended to offer a bridging course prior to exam preparation courses wherein the primary focus is devoted to enhancing candidates' accuracy and communicative competence in discussing unfamiliar topics through corrective feedback. The findings of this study could offer valuable insights to TOEFL iBT test preparation instructors, enabling them to better understand the challenges students face and plan strategies to address these challenges effectively within preparation courses.

1. Introduction

The Test of English as a Foreign Language (TOEFL) has gained unprecedented currency over the past years as it provides a standardized measure of one's proficiency in the English language. Though the test comes in different forms, the internet-based format of the TOEFL test (TOEFL iBT) is probably the most popular one. Many universities use this test as a benchmark to ascertain the eligibility of students, especially regarding exit requirements and university admission (Crystal, 2003; Ng, 2007). Moreover, government agencies and scholarship review boards use TOEFL iBT as part of their eligibility criteria (Antoni, 2014; Herwanda et al., 2012).

On the other hand, merely holding a TOEFL certificate may not meet the stringent criteria of academic institutions. Many of these establishments typically demand a C1 level score, which corresponds to a TOEFL iBT score of 95 and above. Consequently, a substantial number of students and educators dedicate their time and effort to a thorough TOEFL iBT exam preparation. This is because attaining a C1 level score (TOEFL 95 and above) on the TOEFL iBT exam paves the way for exciting opportunities in one's academic and professional journey.

Due to the significance of TOEFL iBT, many students enroll in General English skill development and TOEFL preparation courses to develop the required skills prior to taking the exam. However, despite all the preparation, many candidates find achieving a C1 level score daunting, since

---

[1] Pedagogía de los Idiomas Nacionales y Extranjeros (PINE), Universidad Nacional de Educación (UNAE), Azogues, Cañar, Ecuador. Email: omid.sahranavard@unae.edu.ec

they need to produce advanced-level forms of English (Intriago et al., 2019). While this frustration is evident, its root cause and possible remedies remain largely obscure. Therefore, it is imperative to identify the challenges and barriers that affect obtaining C1 level scores in order to provide exam preparation instructions with pertinent recommendations to tackle them within preparation courses.

Many studies have targeted identifying challenges and difficulties in TOEFL preparation by testing useful exam techniques and examining psychological factors involved in the test-taking process (Akmal et al., 2020; Barnes, 2016; Halim & Ardiningtyas, 2018; Masfufah, 2018; Netta & Trisnawati, 2019). Nevertheless, to the best of the author's knowledge, no study has attempted to align the rubric expectations with student performance to pinpoint the challenges and needs in achieving advanced-level scores (C1 and above). Furthermore, in the broader context of research on TOEFL iBT, it is noteworthy to consider that Ecuadorian candidates have not been studied so far. As highlighted by Bolton (2008), learners in various regions of the world may exhibit distinct characteristics compared to those in other areas. Therefore, when studying high-stakes language assessments, taking the geographic location of test-takers into consideration becomes imperative (Rezaeian et al., 2020).

As Nation and Macalister state (2010), identifying exam preparation needs and challenges may help educators assess their learners' existing knowledge and ascertain their areas of improvement so as to bridge the gap between the desired and actual performance levels. By providing meticulously designed preparatory programs that cater to students' individual needs and align with the examination rubrics, educators can equip learners with the essential language proficiency and test-taking skills to tackle the TOEFL iBT test with confidence. Hence, this study aimed to identify the expectations of the TOEFL iBT exam as per its rubrics and investigate the challenges and needs of Ecuadorian candidates to achieve C1 level scores in the speaking and writing parts of the TOEFL iBT.

2. Review of Literature
2.1. TOEFL iBT

TOEFL was developed by the National Council on the Testing of English as a Foreign Language in 1963 and it is a widely recognized standardized test for assessing the English language proficiency of non-native speakers (Stricker & Attali, 2010). The test is accepted as proof of English language proficiency by over 7,000 academic institutions in 130 countries, including some of the best universities in the world (Rahmah, 2019; Setiawan, 2013; Soali & Pujiani, 2020). The development of the TOEFL speaking and writing parts of the test aligns with an extended version of Bachman's (1990) model of "communicative language" ability and Canale and Swain's (1980) model of "communicative competence", tapping into one's "linguistic knowledge" (i.e., syntactic, textual, and sociolinguistic knowledge), strategic competence, and the contextual use of language (Brooks & Swain, 2014).

According to Taylor & Angelis (2011), the basis for TOEFL iBT was formed in a way that mirrors the fundamental abilities crucial for excelling in English-medium universities. Such skills include the mastery to summarize, combine, and convey important information in lectures and other academic contexts accurately and coherently through the contextual use of language (Biber & Gray, 2013; Enright & Tyson, 2011; Frost et al., 2020). Since the TOEFL iBT test measures candidates' language skills in the academic context, its speaking and writing tasks are centered around topics, subjects, and real problems in academic contexts such as campus problems, lectures, and academic discussions. These tasks are divided into two categories: independent and integrated.

In the independent tasks, candidates are expected to articulate their perspective on an academic subject, whereas in the integrated tasks, they must read academic passages, and listen to lectures and conversations about topics in various fields such as history, astrology, business, and other academic fields in order to succinctly summarize the primary ideas in either spoken or written format. As for the time, candidates only have 45 seconds to respond to the independent question and a mere 60 seconds to respond to the integrated tasks in the speaking part. In the writing part, the time allocated to carry out the writing tasks is 20 minutes for the integrated task and 10 minutes for the independent task.

2.2. Attaining C1 level Scores (95 and above) in TOEFL iBT

TOEFL scores are aligned with the Common European Framework of Reference for Languages (CEFR), which categorizes proficiency levels into six levels: A1, A2, B1, B2, C1, and C2. To be classified as C1, which indicates an advanced level of English proficiency, candidates need to achieve

a minimum score of 24 out of 30 in the writing section and 25 out of 30 in the speaking section (Educational Testing Service, 2021). As per TOEFL iBT rubrics, achieving a C1 level score in the TOEFL speaking and writing sections requires a high level of proficiency and communicative competence. Candidates need to respond to speaking and writing tasks fluently, accurately, and clearly while using advanced forms of language. Additionally, they must display a wide range of vocabulary and precise grammar along with impeccable mechanics in writing (Educational Testing Service, 2021; Educational Testing Service, 2022a, 2022b; Educational Testing Service, 2023a, 2023b).

According to the TOEFL rubrics, errors in speaking and writing production could be deemed acceptable in case they don't hinder understanding. Moreover, candidates must excel in topic development, fully addressing the task and expressing ideas coherently and eloquently. This includes providing well-chosen supporting details and examples, demonstrating logical progression, and impeccable organization. Moreover, candidates are expected to maintain an exceptional focus on the topic throughout their responses by effectively linking ideas and utilizing a wide array of transitions. Furthermore, test takers must employ a variety of delivery strategies such as providing clear reasons, making insightful comparisons, and expressing sophisticated opinions. Candidates would need to convey meaning with precision by utilizing appropriate intonation, stress, and emphasis in speaking and making use of accurate punctuation in writing (Educational Testing Service, 2021; Educational Testing Service, 2022a, 2022b; Educational Testing Service, 2023a, 2023b).

Upon studying TOEFL iBT rubrics, it became evident that the attainment of C1 level scores in TOEFL iBT relies on three fundamental factors: fluency, accuracy, and complexity. According to Choong (2011), fluency pertains to a learner's capability to articulate speech seamlessly and without pauses, while accuracy relates to the learner's capacity to generate language devoid of errors. Conversely, complexity involves the learner's adeptness in employing more advanced linguistic elements, encompassing a diverse array of vocabulary and syntactic structures. It is worth noting that such demand for language proficiency must be exhibited in response to speaking and writing tasks centering around academic topics, which are mostly unfamiliar to candidates. This may hinder achieving the cutoff score of 95.

The literature on the effects of topic familiarity on production reveals that language learners perform better and are able to produce more complex and more accurate language when discussing familiar topics (Biber & Conrad, 2009; Lee & Anderson, 2007; Roohani et al., 2017; Skehan, 2009; Tedick, 1988). On a different note, a collective body of research highlights the significance of feedback in promoting improved performance corrective feedback plays a crucial role in closing the gap between the actual performance and the desired one (Ajabshir & Ebadi, (2023); Cheng et al., 2021; Zhang & Zhang, 2022).

2.3. Challenges and Needs in TOEFL iBT Test Preparation

While examining the TOEFL iBT rubrics highlights the significance of fluency, accuracy, and the use of sophisticated vocabulary and grammar, most prior studies on TOEFL iBT preparation courses have investigated applicants' and teachers' perceptions of the test structure, psychological factors, and the usefulness of the preparation courses and materials. It appears that identifying the challenges and needs of students with reference to the expectations outlined in TOEFL iBT rubrics has largely remained unexplored.

In a study, Alderson and Hamp-Lyons (1996) aimed to identify differences in teaching methods between TOEFL preparation classes and General English classes by analyzing factors such as turn-taking, test-taking time, teacher talk, laughter frequency, and pair work. The findings revealed that TOEFL preparation classes placed more emphasis on test-taking strategies and had fewer interactive activities compared to the regular ESL classes. In this instance, the quality of teaching relied on the attitudes and convictions of instructors (Alderson & Hamp-Lyons, 1996; Cheng & Watanabe, 2004). Naturally, focusing solely on test preparation without a balanced approach can yield unfavorable outcomes. For instance, Takanashi's research in 2004 uncovered that Japanese students encounter difficulties in effectively using communicative English and often struggle to attain their desired TOEFL scores, despite their strong enthusiasm for learning the English language in general. Moreover, as Spratt (2005) pointed out, the structure of the test significantly influences teaching methods, a phenomenon

often referred to as "washback." This highlights the reciprocal relationship between test design and how English is taught.

The academic nature of topics discussed in the TOEFL iBT exam is another factor affecting teaching, learning, and learners' performance. Huang's (2006) evaluation of ESL students from North American campuses revealed that Chinese students, despite achieving high TOEFL scores, faced difficulties comprehending academic lectures and delivering presentations. This underscores the need for proficiency beyond mere test scores. Similarly, Barnes (2016) found that the majority of TOEFL teachers primarily focused on delivering instructions and facilitating individual practice for students and that the TOEFL classes provide fewer opportunities for pair or group work when compared to general English classes.

In a similar vein, Mahmud (2014) determined that many test takers tended to respond to TOEFL writing tasks by relying on memorized patterns, to compensate for their shortcomings. To combat this, Pan and In'nami (2017) argue that educators should allocate a greater portion of their instructional time to activities that involve students in both test preparation and the acquisition of English skills essential for professional contexts.

Halim and Ardiningtyas (2018) conducted research on the difficulties encountered by students when responding to TOEFL test questions. Their findings indicated that a substantial number of students struggled to perform well on the test due to lower language proficiency, insufficient practice, and limited motivation. Furthermore, Masfufah (2018) found that Malaysian students primarily enrolled in preparation programs due to their inadequate English proficiency. Importantly, these issues are not exclusive to English learners, as they are also prevalent among English major students who have already attained language proficiency (Akmal et al., 2020).

In the broader context of high-stakes assessment, Mohammadi et al. (2023) discovered that teacher-only and blended feedback modes lead to a notable improvement in candidates' writing, specifically lexical resources. Furthermore, the participants' attitudes toward feedback modes were highly positive. In a different study, Arefsadr et al. (2022) investigated the reasons behind the consistently low scores of IELTS candidates in writing by comparing candidates' performance in IELTS and World Englishes (WEs) essay writing tasks. The results revealed that candidates scored better in the WEs test due to the WEs's structure, which provides more time for writing, clearer instructions, and more authentic writing tasks. In light of the reviewed literature, the following research questions were formulated:

1. What factors prevent candidates from achieving C1 level scores in the speaking and writing parts of the TOEFL iBT exam?
2. What are TOEFL iBT candidates' language needs to achieve C1 level scores in speaking and writing according to the TOEFL iBT rubrics?

3. Method
3.1. Research Design

The present study aimed to investigate the challenges and needs of TOEFL candidates striving to attain C1 level scores in the writing and speaking sections of a TOEFL iBT preparation course. To accomplish this objective, a mixed-method research approach was employed. Specifically, this study adopted the Sequential Explanatory Strategy to collect quantitative and qualitative data. According to Creswell and Creswell (2017), the Sequential Explanatory Strategy is a research approach characterized by two distinct phases. Initially, quantitative data is collected, then, qualitative data is gathered to enhance the understanding and interpretation of quantitative results. Therefore, the researcher gathered quantitative data through a survey in the initial phase of the study, then collected qualitative data through class observations and the analysis of TOEFL iBT mock tests. This integration of data allowed for the enrichment of quantitative findings by the subsequent qualitative data collection.

3.2. Participants

The participants of this study were chosen from two intact TOEFL iBT preparation classes, totaling 46 students (24 in one class and 22 in the other), at Universidad Nacional de Educación in Ecuador. The composition of the participants comprised 12 male students (26%) and 34 female students (74%). Their ages ranged from 22 to 30 years, with an average age of 24 and a standard deviation of

2.71. It is noteworthy that all participants in this study had successfully completed a four-year General English program at the same university, progressing from an A1 level in the first semester to a C1 level by the eighth semester. In this preparation course, the participants received two months of instruction focusing on exam technique, scoring criteria, rubrics, and various activities to develop speaking and writing skills. The participants frequently took TOEFL iBT mock exams to ensure optimal preparation and familiarity with the exam structure and their scores upon taking the exam. Given the circumstances, the researcher functioned as the teacher of this preparation course.

3.3. Data Collection

Since this study followed a mixed-method approach to data collection, the data were gathered in two types: qualitative and quantitative. Moreover, since the population under study was not large and all participants possessed similar characteristics, sampling was not done and the study was conducted on the whole population (n = 46).

3.3.1. Quantitative Data Collection

To assess the students' self-perceived performance, a survey questionnaire was administered. This questionnaire, designed by the researcher, consisted of three distinct sections aimed at evaluating language proficiency, identifying challenges associated with the exam, and uncovering the students' specific needs to succeed in the TOEFL iBT exam. As evident in Table 1, the reliability of the questionnaire was measured by Cronbach's Alpha. The results of the reliability analysis yielded a coefficient of .88, indicating a commendable level of reliability.

Table 1

Reliability Statistics

Cronbach's Alpha	N of Items
.883	37

3.3.2. Qualitative Data Collection

In addition to the survey, qualitative data were gathered by recording class observations in a teaching journal. In this sense, the teacher closely monitored the students' performance throughout the two-month instruction period, paying particular attention to fluency, accuracy, and the use of a range of words and grammar as per the rubrics. Then, the researcher interpreted the results by matching the observed data to the expectations outlined in the TOEFL iBT rubrics to identify disparities between the expectations of the test and candidates' performance.

Furthermore, the students participated in a final mock test, which served as another source of qualitative data. The researcher analyzed the results of this mock test using the same TOEFL rubrics, employing a qualitative approach once again. The criteria for data analysis in this part of the study also focused on evaluating the students' fluency, accuracy, range of words, and grammar, as outlined in the TOEFL iBT rubrics.

4. Results
4.1. Quantitative Results

As presented in Table 2, the survey results showed that only a small percentage of students (2.1%) considered themselves capable of producing responses at the C1 level, and the majority (87.3%) lacked confidence in their English proficiency.

Table 2

*Students' Level of Proficiency According to CEFR (Survey)*

CEFR Level	N	%
B1	21	43.17%
B2	25	53.2%
C1	1	2.1%

In order to identify the reasons behind the participants' lack of confidence in their proficiency, the second part of the questionnaire was analyzed. The results showed that students' fluency, accuracy, and ability to develop ideas using a range of vocabulary and grammar were among the specific reasons hindering the participants' proficiency. As per Figure 1, only a small percentage of students (6.4%) believed they were fluent enough to achieve C1 level scores in speaking and an even lower percentage (2.1%) felt accurate in their speaking.

Additionally, very few students (4.3%) believed they could develop their ideas using a range of grammar in speaking, and none (0%) felt they could use a range of vocabulary effectively for the same purpose. The numbers in writing were similarly low. Only 2.1% of students reported C1 level accuracy, and 12.8% felt they could develop ideas at a C1 level. The use of a range of vocabulary and grammar in writing was also lacking, with only 8.5% and 4.3% of students, respectively, believing they could effectively employ these skills.

Figure 1

*Specific Challenges According to the TOEFL iBT Rubric (Survey)*
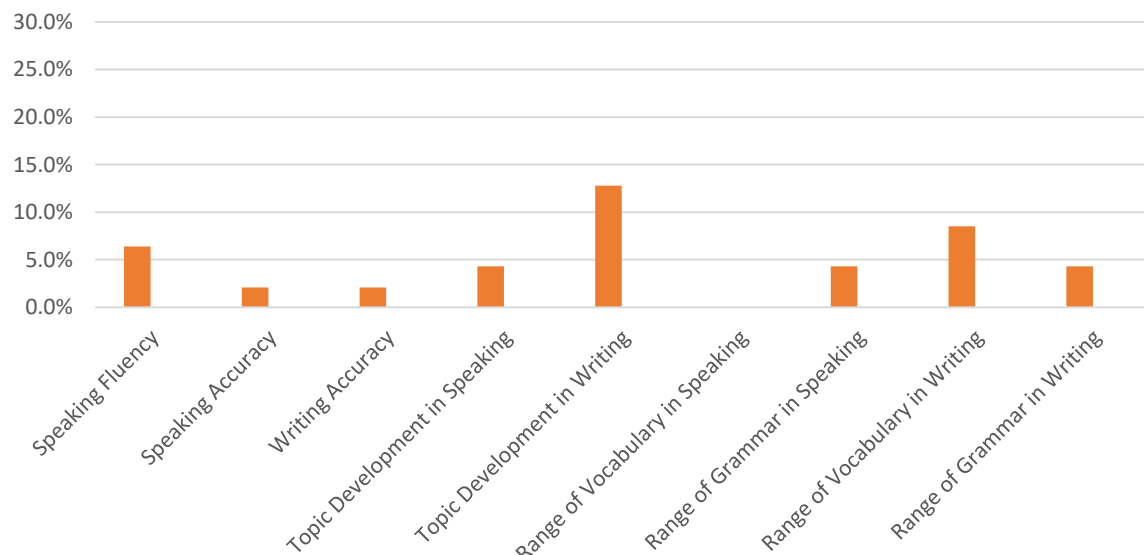
The third part of the survey sought to identify whether or not having more thinking time would affect students' performance. This is a significant factor because it provides valuable insight into students' fluency and automaticity, a determining factor in obtaining C1 level scores at the TOEFL iBT exam. The findings revealed that students strongly believed that having more time to think and prepare their answers would significantly enhance their TOEFL performance and scores. As seen in Figure 2, More than 60% of participants believed having more thinking time would positively affect their fluency, and 61.7% reported a positive impact on their accuracy. Adequate thinking time was also seen as beneficial for vocabulary and grammar.

Nearly half of the students (48.1%) expressed that thinking time had a high impact on effectively using a wide vocabulary repertoire and 48.9% of students felt that thinking time greatly

improved their grammar usage. Interestingly enough, no student thought additional thinking time to be unhelpful in providing better responses. The fact that students believed they could perform better with more preparation time proves that they had yet to develop the required automaticity in responding to TOEFL iBT tasks and needed more time to think and plan their answers.

Figure 2

*The Effect of Having More Thinking Time on Performance (Survey)*
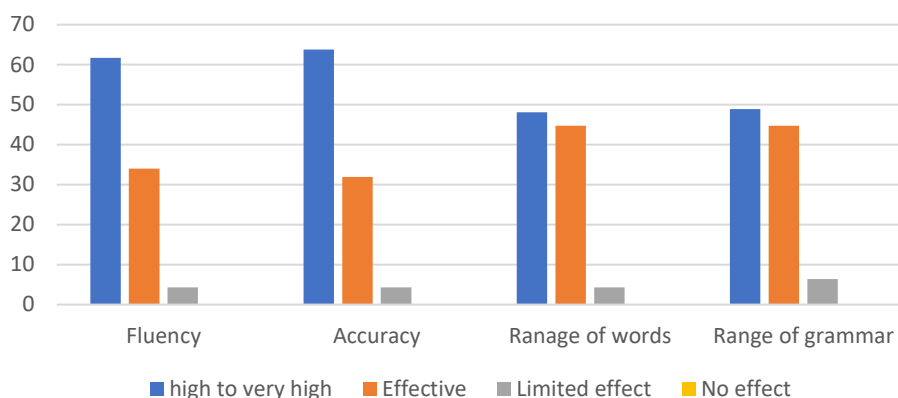
The last part of the survey focused on students' needs. This part explored various aspects of language improvement, including general speaking and writing skills, as well as specific areas such as fluency, accuracy, topic development, range of grammar, vocabulary, and pronunciation. The results showed that students recognized the need to practice exam techniques more and improve their overall speaking and writing skills, as well as, fluency, accuracy, and the use of sophisticated vocabulary and grammar forms.

As demonstrated in Figure 3, the highest priority among students was exam technique, with 74.4% expressing a need for more practice in this area. Fluency and accuracy in speaking were also significant needs, with 74.5% and 70.2% of students indicating a high to very high need, respectively. Using a range of vocabulary and grammar in both speaking and writing was also a notable need with 68.1 and 70.2 percent of students believing they needed more improvement in the use of grammar and vocabulary respectively. Pronunciation was another important area, with nearly 60% of students desiring improvement in order to achieve better scores on the TOEFL iBT exam. In writing, accuracy was identified as the top priority, with 70.2% of students indicating a high to very high need for precise and correct language usage. Effective topic development was also important, as 63.8% of students expressed a high to very high need in this aspect.

Figure 3

*Specific Needs as Reported by Students (Survey)*
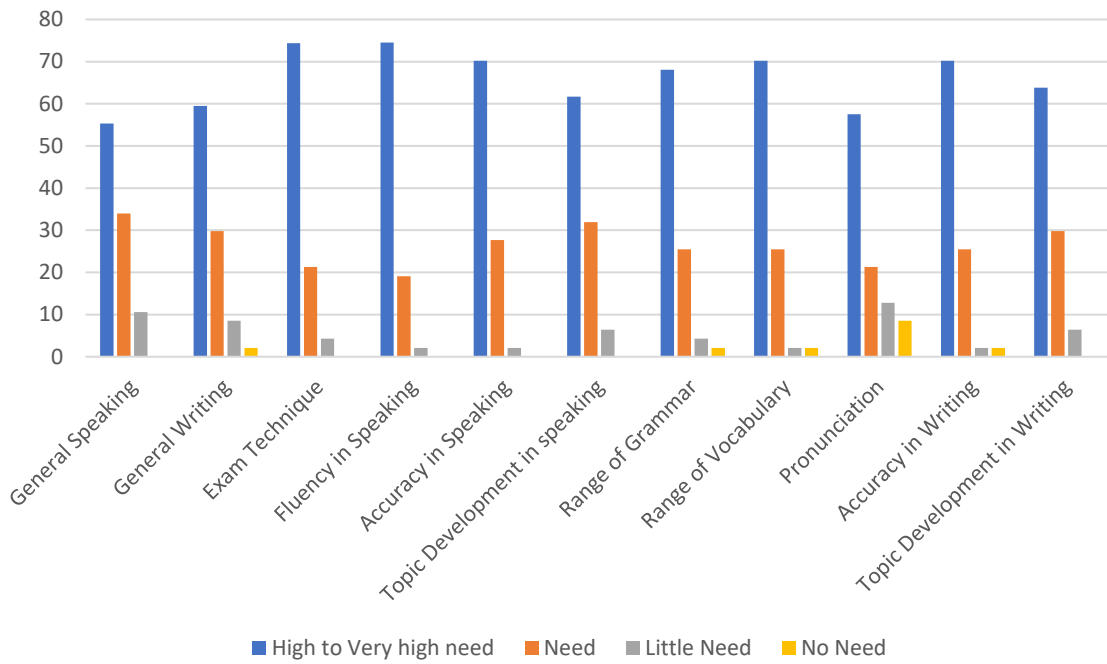
4.2. Qualitative Results

The teacher's observations and mock test results further supported and enriched the quantitative findings. From the initial stages of observation, it became evident to the researcher that students encountered difficulties in providing fluent responses, and that their proficiency level appeared insufficient. As seen in figures 4 and 5, a low percentage of participants (4.26%) achieved a C1 score in speaking in the mock test, whereas in writing, a higher percentage (29.79%) could produce C1 level drafts.
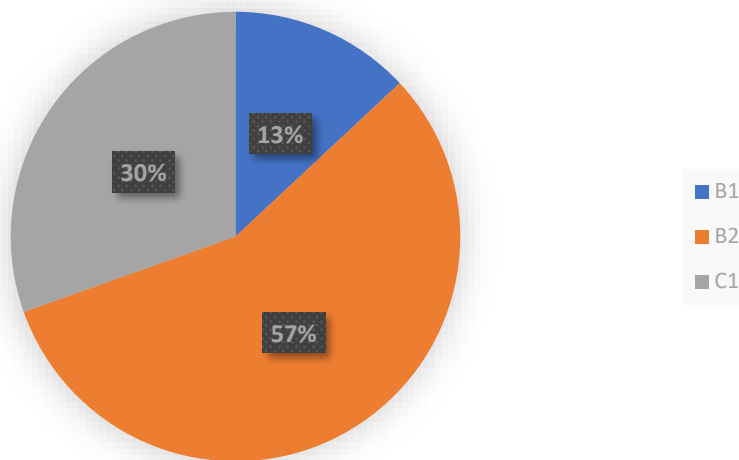
Figure 4

*Students' CEFR Levels in Speaking (Mock Exam)*
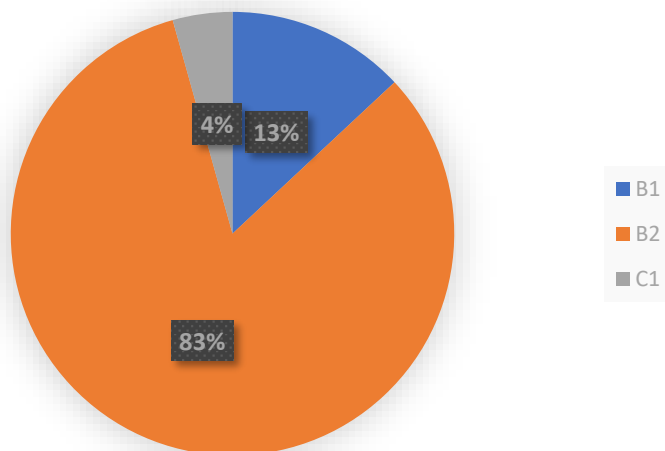
Figure 5

Students' CEFR Levels in Writing (Mock Exam)

Though the observations did not paint a detailed picture, the analysis of the mock exams shed light on the observation data. In conducting a thorough examination of the students' grades in the final TOEFL iBT exam, a detailed analysis revealed notable disparities in the performance of students in speaking and writing. As Figure 6 demonstrates, the average performance of the students in the TOEFL speaking section was found to be 16.70% lower than the expected proficiency level denoted as C level. Similarly, in the TOEFL writing section, the students exhibited an average performance that was 11.87% lower than the C-level. These findings indicate a deviation from the anticipated levels of proficiency in both speaking and writing skills among the student population under study.

Figure 6

Average Distant from C1 Level (Mock Exams)
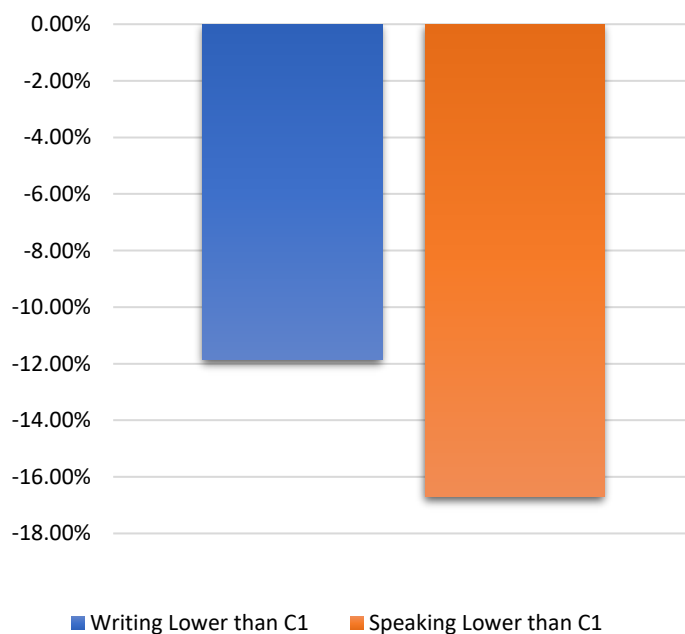
Moreover, the teacher's observations in the later stages of monitoring students' performance supported the survey findings, highlighting issues with fluency, accuracy, grammar, word choice, and topic development among students. The observations also noted the influence of students' first language

(L1) as the main factor influencing inaccurate production. On a different note, students often struggled to provide C1 level responses to questions on academic topics, which were unfamiliar to them, but performed well and were confident while discussing general and familiar topics. However, with additional time to think and organize their thoughts, students were able to express themselves more effectively and coherently, leading to increased fluency and reduced errors. Furthermore, the extra thinking time allowed students to utilize a wider range of vocabulary and demonstrate a greater variety of grammatical structures in their spoken and written communication.

This disparity in performance indicates that students need to spend more time practicing oral and written forms of language on academic topics. Furthermore, even though the teacher reviewed C1 level grammar and vocabulary with the students and provided them with lists of words and grammar to use in their speaking and writing, they failed to incorporate the suggested words and grammar into their responses. This suggests that the students' communicative competence was lower than their linguistic competence, hindering their ability to reach C1 level scores in speaking and writing.

As per Figures 7 and 8, the mock exams further confirmed the lack of fluency, and accuracy, and exhibited the use of simple language and memorized phrases instead of original communication. In the speaking section, the most challenging area for students was using a range of words and grammar, which made up 10.85% of the total 16.70% disparity from C1 level speaking. This was followed by fluency (2.00%), and accuracy (3.01%). Within the writing section, the most challenging area for students was once again the range of words and grammar, which made up 7.71% of the total 11.87% performance deficit. Accuracy was another significant challenge accounting for 2.14% of the overall disparity from C1 level production in writing.

Figure 7

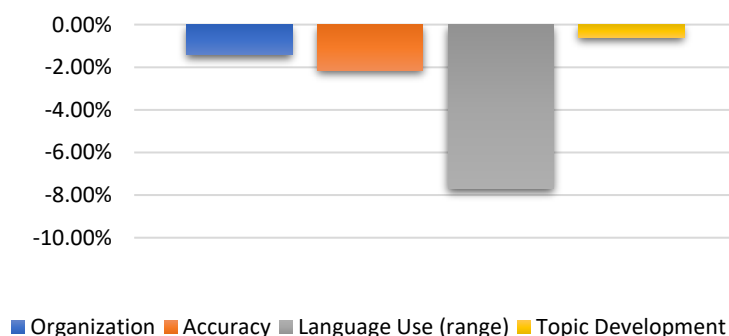*Below C1 Level Scores in Speaking (Mock Exam)*

■ Fluency ■ Accuracy ■ Language Use (range) ■ Topic Development
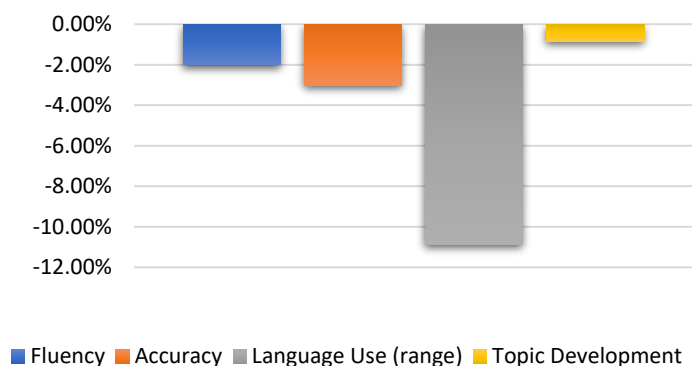
Figure 8

Below C1 Level Scores in Writing (Mock Exam)

■ Organization ■ Accuracy ■ Language Use (range) ■ Topic Development

5. Discussion

The results of the test, observation, and survey revealed that the majority of students did not possess C1 level proficiency (according to TOEFL iBT rubrics) as they struggled to use a range of grammar and vocabulary, and showed subpar fluency and accuracy in their production. Consequently, they found speaking and writing tasks burdensome and required more time to plan their answers to make up for their shortcomings. Another challenge was the fact that students had the tendency to use simple vocabulary and grammar structures. Lastly, the presence of unfamiliar academic topics in the speaking and writing tasks drastically increased the difficulty of the exam for the candidates.

From the aforementioned challenges, the use of simple vocabulary and grammar can be considered as part of the students' "interlanguage" journey. As per the characteristics of interlanguage (Selinker, 1972), learners often tend to use simple and familiar words and sentence patterns from their native language, while gradually incorporating elements of the target language as they progress in their language-learning journey. Additionally, as the reviewed literature (Biber & Conrad, 2009; Lee & Anderson, 2007; Roohani et al., 2017; Skehan, 2009; Tedick, 1988) revealed, students tend to use more complex language in discussing familiar topics. This holds significance since the topics discussed in the TOEFL exam are academic in nature and generally unfamiliar to students.

Moreover, this study highlighted the influence of L1 as the dominant factor that hindered students' accuracy, even after devoting a significant amount of time to corrective feedback to tackle inaccurate oral and written production. This signals the fossilization of inaccurate stretches of language influenced by L1. To combat this, the reviewed literature suggests the use of feedback corrective feedback (Ajabshir & Ebadi, 2023; Cheng et al., 2021; Zhang & Zhang, 2022). Implementing nuanced corrective feedback techniques that address errors and inaccurate utterances using both on-the-spot and delayed corrective feedback methods could be a viable tool to help students overcome their mistakes.

Nevertheless, it should be noted that the participants' average speaking proficiency was just 16.70% below the C1 level, and their writing proficiency fell short by a mere 11.87%. Specifically, the challenges they encountered in terms of fluency, accuracy, and the utilization of a diverse vocabulary and grammar were found to be only 2.00%, 3.01%, and 7.71% lower than the C1 level, respectively. With such a narrow margin separating their current proficiency from the desired C1 level, it becomes evident that more time and practice are needed to bridge this gap successfully. As Larsen-Freeman (2012) notes, developing language skills is a gradual and time-intensive process involving interaction and repeated error correction in various contexts.

Considering the gradual nature of language learning and the close proximity of the participants' proficiency to the C1 benchmark, it could be concluded that students might have been able to attain C1 level scores with additional training and time to practice productive skills. It appears that the participants of this study simply needed more time to practice their target language in discussing academic and unfamiliar topics and improve their level of proficiency to advance in their "interlanguage" journey. Moreover, the majority of the participants indicated that they needed to improve their fluency, accuracy, and complexity in responding to the TOEFL iBT tasks. Therefore, it could be concluded that the needs of the participants lie in providing them with more opportunities to practice the target language in discussing academic and unfamiliar topics could enhance candidates' fluency and automaticity.

Hence, the researcher proposes introducing a bridging course that provides ample opportunity for oral and written practice about unfamiliar and academic topics and incorporates corrective feedback guidelines before starting exam preparation. Such a course would not only offer abundant opportunities for students to hone their productive skills on various topics, but also help them remedy their L1-influenced errors through repeated error-correction activities. Moreover, a bridging course could provide students with sufficient thinking and planning time to utilize a range of vocabulary and grammatical structures in their production.

Utilizing a bridging course offers a less stressful environment compared to preparation courses, allowing students to improve their proficiency without undue pressure. By fostering a supportive and welcoming atmosphere, students feel comfortable responding to challenging questions and expressing their opinions without time constraints. This nurturing environment, coupled with the provision of corrective feedback can assist students to enhance their fluency, accuracy, topic development, and the utilization of a broader vocabulary and grammatical repertoire effectively.

Finally, the conclusions drawn from this study are in line with the results of the studies done by Halim and Ardiningtyas (2018), as well as Takanashi (2004) who attributed shortcomings in TOEFL preparation to relatively low level of language proficiency and lack of practice. Also, this study corroborates the call made by Pan and In'nami (2017) regarding the essential nature of involving students in the acquisition of English skills when preparing for exams. Additionally, similar to the study conducted by Huang (2006) the students in this study also experienced difficulties in responding to tasks involving academic lectures, despite their proficiency in general English. Finally, the participants in this study had a tendency to memorize complex patterns of language and uncommon vocabulary as a remedy to overcome their substandard proficiency, which is in line with the results of the study conducted by Mahmud (2014).

6. Conclusion

In light of the findings of the study, some pedagogical implications for pertinent stakeholders are offered. Firstly, this study sheds light on some challenging areas in the context of TOEFL exam preparation that need to be focused on. This study may lead teachers to adjust their instructional methods in order to tackle the specific challenges students face during exam preparation programs. For example, they can incorporate tailored activities and exercises aimed at bolstering grammar and vocabulary proficiency, fostering greater fluency, and enhancing the accuracy of students' speaking and writing tasks. Acknowledging the potential hurdles in achieving C1 level proficiency is equally critical. By setting realistic goals and expectations, and modifying the curriculum to meet those goals, educators can guide their students in a more constructive manner, thus diminishing the anxiety and frustration often associated with high-stakes language assessments. Lastly, as identified by this study, the implementation of continuous assessment and feedback mechanisms is crucial. Exam preparation instructors can put more time and attention into monitoring candidates' progress and provide timely guidance for improvement, offering a particularly effective means of addressing the specific issues identified in vocabulary, grammar, fluency, and accuracy.

Nevertheless, the present study has some limitations which need to be considered. One significant limitation of this study lies in the constrained sample size employed by the researcher. The use of a larger sample could have enhanced the study's external validity, allowing for more generalized conclusions about the population under investigation. Additionally, the absence of multiple raters for assessing students' speaking and writing performance, which led to not considering the interrater reliability measures concerning the scores, introduced a potential source of bias and reliability concerns. Having multiple raters and assessing interrater reliability would have strengthened the study's methodological rigor, ensuring greater confidence in the validity of the findings and reducing the susceptibility to individual rater biases. Future research could aim to employ a larger and more diverse sample to enhance the generalizability of findings and gain a more comprehensive understanding of the studied phenomena.

Lastly, the findings of this study highlight a few areas for future research. Firstly, identifying common L1-influenced speaking and writing errors in L2 learners of English could help educators consider and prepare appropriate remedial plans. Therefore, it appears that targeting the identification of common mistakes and errors in future studies could provide significant value to instructors who teach exam preparation courses. Secondly, it seems that fossilization plays a crucial role in students' erroneous productions. Hence, investigating nuanced remedial plans that help students overcome fossilized inaccuracies seems important. Another area of investigation could refer to the effects of incorporating more academic and unfamiliar topics into C1 level general courses. Analyzing the impact of discussing TOEFL-like academic and unfamiliar topics on learners' language proficiency and their motivation could yield valuable insights for educators aiming to enhance their students' language abilities in these domains.

Declaration of Conflicting Interests
The author has affirmed that there are no conceivable conflicts of interest regarding the research, authorship, and publication of this article.

Funding

References

Ajabshir, Z. F., & Ebadi, S. (2023). The effects of automatic writing evaluation and teacher-focused feedback on CALF measures and overall quality of L2 writing across different genres. *Asian-Pacific Journal of Second and Foreign Language Education*, *8*(1), 26, 1-16. https://link.springer.com/article/10.1186/s40862-023-00201-9

Akmal, S., Risdaneva, R., Habiburrahim, H., & Sari, M. (2020). The English teachers' challenges in TOEFL preparation for senior high school students. *Journal on English as a Foreign Language, 10*(1), 25-45. https://doi.org/10.23971/jefl.v10i1.1627

Alderson, J. C., & Hamp-Lyons, L. (1996). TOEFL preparation courses: A study of washback. *Language Testing*, *13*(3), 280–297. https://doi.org/10.1177/026553229601300304

Antoni, R. (2014). An analysis on 6th semester students' TOEFL experience at English department of teachers training and education faculty of Pasir Pengaraian University. *Edu Research*, *3*(1), 9-16. https://media.neliti.com/media/publications/59684-EN-an-analysis-on-6th-semester-students-toe.pdf

Arefsadr, S., Babaii, E., & Hashemi, M. R. (2022). Why IELTS candidates score low in writing: Investigating the effects of test design and scoring criteria on test-takers' grades in IELTS and World Englishes essay writing tests. *International Journal of Language Testing*, *12*(2), 145-159. https://doi.org/10.22034/ijlt.2022.157131

Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford University Press.

Barnes, M. (2016). The washback of the TOEFL iBT in Vietnam. *Australian Journal of Teacher Education*, 41(7), 158–174. https://doi.org/10.14221/ajte.2016v41n7.10

Biber, D., & Conrad, S. (2009). *Register, genre, and style*. Cambridge University Press.

Biber, D., & Gray, B. (2013). Discourse characteristics of writing and speaking task types on the TOEFL ibt® test: A lexico-grammatical analysis. *ETS Research Report Series*, *2013*(1), i-128. https://doi.org/10.1002/j.2333-8504.2013.tb02311.x

Bolton, K. (2008). English in Asia, Asian Englishes, and the issue of proficiency. *English Today, 24*(2), 3-12. http://dx.doi.org.ezproxy.cul.columbia.edu/10.1017/S026607840800014X

Brooks, L., & Swain, M. (2014). Contextualizing performances: Comparing performances during TOEFL iBTTM and real-life academic speaking activities. *Language Assessment Quarterly*, *11*(4), 353-373. https://doi.org/10.1080/15434303.2014.947532

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, *1*(1), 1-47. http://www.uefap.com/tefsp/bibliog/canale_swain.pdf

Cheng, L., & Watanabe, Y. (Eds.). (2004). *Washback in language testing: Research contexts and methods*. Routledge.

Cheng, X., Zhang, L. J., & Yan, Q. (2021). Exploring teacher written feedback in EFL writing classrooms: Beliefs and practices in interaction. *Language Teaching Research, 0*(0). https://doi.org/10.1177/13621688211057665

Choong, K. P. (2011). Task complexity and linguistic complexity: An exploratory study. *Studies in Applied Linguistics and TESOL*, *11*(1). https://doi.org/10.7916/salt.v11i1.1413

Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.

Crystal, D. (2003). *English as a global language*. Cambridge University Press.

Educational Testing Service. (2021). *Performance descriptors for the TOEFL iBT test*. Educational Testing Service. https://www.ets.org/pdfs/toefl/toefl-ibt-performance-descriptors.pdf

Educational Testing Service. (2022a). *TOEFL iBT independent speaking rubrics*. Educational Testing Service. https://www.ets.org/pdfs/toefl/toefl-ibt-speaking-rubrics.pdf

Educational Testing Service. (2022b). *TOEFL iBT integrated speaking rubrics*. Educational Testing Service. https://www.ets.org/pdfs/toefl/toefl-ibt-speaking-rubrics.pdf

Educational Testing Service. (2023a). *TOEFL iBT integrated writing rubric*. Educational Testing Service. https://www.ets.org/pdfs/toefl/toefl-ibt-writing-rubrics-enhanced.pdf

Educational Testing Service. (2023b). *TOEFL iBT writing for an academic discussion Rubrics*. Educational Testing Service. https://www.toefl-ibt.jp/dcms_media/other/toeflibt_writing_for_an_academic_discussion_rubrics.pdf

Enright, M., & Tyson, E. (2011). Validity evidence supporting the interpretation and use of TOEFL iBT scores. *TOEFL iBT Research Insight 4*(1)*,* 1-21. https://www.ets.org/pdfs/toefl/toefl-ibt-insight-s1v4.pdf

Frost, K., Clothier, J., Huisman, A., & Wigglesworth, G. (2020). Responding to a TOEFL iBT integrated speaking task: Mapping task demands and test-takers' use of stimulus content. *Language Testing*, 37(1), 133–155. https://doi.org/10.1177/0265532219860750

Halim, N., & Ardiningtyas, S. Y. (2018). Difficulties faced by the students in answering TOEFL test questions. *English, Teaching, Learning, and Research Journal*, 4(2), 219-231. https://doi.org/10.24252/eternal.v42.2018.a7

Herwanda, R., Safriyono, DA, & Haryono, PY (2012). Evaluation of the TOEFL matriculation program for Al Azhar university students in Indonesia 2010/2011. *Journal of Al Azhar Indonesia Seri Humaniora*, *1*(3), 179-188. https://jurnal.uai.ac.id/index.php/SH/article/download/67/70

Huang, J. (2006). English abilities for academic listening: How confident are Chinese students? *College Student Journal*, *40*(1), 218-226. https://link.gale.com/apps/doc/A145156547/AONE?u=sri&sid=googleScholar&xid=7e4e3347

Hyland, K. & Hyland, F. (2006). *Feedback in second language writing: Contexts and issues*. Cambridge University Press.

Intriago, W., Intriago, G., & Hidalgo, D. (2019). Uso de EMI en el sistema educativo ecuatoriano, una reflexión objetiva. *Opuntia Brava*, *11*(1), 97-107. https://opuntiabrava.ult.edu.cu/index.php/opuntiabrava/article/view/700

Larsen-Freeman, D. (2012). From unity to diversity: Twenty-five years of language-teaching methodology. In *English Teaching Forum, 50*(2), 28-38. https://eric.ed.gov/?id=EJ982846

Lee, H. K., & Anderson, C. (2007). Validity and topic generality of a writing performance test. *Language Testing, 24*(3), 307-330. https://doi.org/10.1177/0265532207077200

Mahmud, M. (2014). The EFL students' problems in answering the test of English as a foreign language (TOEFL): A study in Indonesian context. *Theory and Practice in Language Studies*, *4*(12), 2581-2587. http://dx.doi.org/10.4304/tpls.4.12.2581

Masfufah, S. (2018). Indonesian college students' perceptions of TOEFL preparation class. *EduLite: Journal of English Education, Literature and Culture, 3*(1), 66-78. http://dx.doi.org/10.30659/e.3.1.66-78

Mohammadi, M., Zarrabi, M., & Kamali, J. (2023). Formative assessment feedback to enhance the writing performance of Iranian IELTS candidates: Blending teacher and automated writing evaluation. *International Journal of Language Testing, 13*(1). https://doi.org/10.22034/ijlt.2022.364072.1201

Nation, I. S. P., & Macalister, J. (2010). *Language curriculum design*. Routledge.

Netta, A., & Trisnawati, I. K. (2019). ACEHNESE undergraduate students' strategies in preparing for TOEFL prediction: A preliminary study. *Englisia*, 7(1), 41-52. http://dx.doi.org/10.22373/ej.v7i1.5779

Ng, J. (2007). *Test of English as a foreign language (TOEFL): Good indicator for students' success at community colleges?* Oregon State University, Oregon.

Pan, Y. C., & In'nami, Y. (2017). Does TOEIC as a University Exit Test Ensure Higher Employability in Taiwan?. *International Journal of Language Testing*, 7(1), 1-27. https://www.ijlt.ir/article_114435.html

Rahmah, S. (2019). *The students' problems in listening comprehension of TOEFL test* (Bachelor's thesis, Universitas Islam Negeri Ar-Raniry Banda Aceh). Institutional Repository of UIN Ar-Raniry. https://repository.ar-raniry.ac.id/id/eprint/9648/1/OKE.pdf

Rezaeian, M., Seyyedrezaei, S.H., Barani, G., & Seyyedrezaei, Z.S. (2020). Construction and validation of educational, social and psychological consequences questionnaires of EPT as a high-stakes

test. *International Journal of Language Testing, 10*(2), 33–70. https://www.ijlt.ir/article_118014.html

Roohani, A., Dayeri, K., & Farhang-Ju, M. (2017). Role of Background Knowledge in Iranian EFL Learners' Reading Comprehension Test Performance. *International Journal of Language Testing*, 7(1), 28-39. https://www.ijlt.ir/article_114436.html

Rosenfeld, M., Leung, S., & Oltman, P. K. (2001*). The reading, writing, speaking, and listening tasks important for academic success at the undergraduate and graduate levels*. Educational Testing Service.

Selinker, L (1972). "Interlanguage". *International Review of Applied Linguistics*. *10* (1–4): 209–241. https://doi.org/10.1515/iral.1972.10.1-4.209

Setiawan, A. (2013). *A study on difficulties of listening comprehension of TOEFL at state islamic college of Jurai Siwo Metro* (Bachelor's thesis, IAIN Metro). AIN Metro Digital Repository. https://repository.metrouniv.ac.id/id/eprint/3052/

Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics, 30*(4), 510–532. https://doi.org/10.1093/applin/amp047

Soali, M., & Pujiani, T. (2020). An error analysis of non- English students department in answering listening section in TOEFL. *J-LEE - Journal of Law, English, and Economics, 2*(01), 21–39. https://doi.org/10.35960/j-lee.v2i01.593

Spratt, M. (2005). Washback and the classroom: The implications for teaching and learning of studies of washback from exams. *Language Teaching Research, 9*(1), 5-29. https://doi.org/10.1191/1362168805lr152oa

Stricker, L. J., & Attali, Y. (2010). Test takers' attitudes about the TOEFL iBTTM . *ETS Research Report Series, 2010*(1). i-16. http://dx.doi.org/10.1002/j.2333-8504.2010.tb02209.x

Takanashi, Y. (2004). *TEFL and communication styles in Japanese culture. Journal of Language Culture and Curriculum, 17* (1). 1-14, https://doi.org/10.1080/07908310408666678

Taylor, C. A., & Angelis, P. (2011). The evolution of the TOEFL. In Chapelle, C. A., Enright, M. K., & Jamieson, J. M. (Eds.). *Building a validity argument for the Test of English as a Foreign Language* (pp. 27-54). Routledge.

Tedick, D. J. (1988). *The effects of topic familiarity on the writing performance of nonnative writers of English at the graduate level* (Doctoral dissertation). The Ohio State University. Proquest. https://www.proquest.com/openview/661b52fd6f6732535ed836e4fcd69de5/1?pq-origsite=gscholar&cbl=18750&diss=y

Zhang, J., & Zhang, L. J. (2022). The effect of feedback on metacognitive strategy use in EFL writing. *Computer Assisted Language Learning*. https://doi.org/10.1080/09588221.2022.2069822