



JEELS

(Journal of English Education and Linguistics Studies)

P-ISSN: 2407-2575 E-ISSN: 2503-2194

<https://jurnalfaktarbiyah.iainkediri.ac.id/index.php/jeels>

THE EFFECTIVENESS OF TOEFL PROGRAM ON THE STUDENTS' ENGLISH PROFICIENCY: A CASE OF AN INDONESIAN UNIVERSITY

*Ahmad Amin Dalimunte¹; Chanata Somanawattana²; Deasy Yunita Siregar³

^{1,3}English Language Education, Universitas Islam Negeri Sumatera Utara, Medan, Indonesia; ²Chandrakasem Rajabhat University, Bangkok, Thailand

*amindalimunte@uinsu.ac.id; chanata.s@chandra.ac.th;

deasyyunitasiregar@uinsu.ac.id

(*) Corresponding Author

Abstract: Having sufficient English language ability is considered important for university students to begin their undergraduate studies. English serves as one of the principal working languages in ASEAN nations, necessitating that ASEAN students complete an English test from graduation. This mandate would positively influence education, tourism, commerce, and investment in ASEAN regions. The present study was to measure the effect of the TOEFL preparation program on English language proficiency of Islamic economics and business students in Indonesia. Using one group pre-test and post-test design, an experimental group consisting of 20

¹Citation in APA style:

Dalimunte, A. A., Somanawattana, C., & Siregar, D. Y. (2025). The effectiveness of TOEFL program on the students' English proficiency performance: A case of an Indonesian university. *JEELS*, 12(1), 141-173. DOI: 10.30762/jeels.v12i1.3970

Submission: November 2024, Revision: Januari 2025, Publication: March 2025

students was randomly selected from 50 students majoring in various Islamic economic and business programs in the first year. The experimental group was intensively taught and learned TOEFL preparation materials using a wide range of teaching and learning strategies and activities. Their ability was measured prior to and after the teaching intervention. The results of statistical measurement using t-test indicate a significant difference between the students mean score of both the pre-test and the post-test. Thus, this one group pre-test and post-test study served as an experiment to enhance students' English language proficiency particularly their listening, reading, and grammar mastery respectively. The findings of the present study provide new insights for teachers and educational institutions in developing effective TOEFL preparation programs to improve students' English skills in various educational institutions.

Keywords: *English language proficiency, grammar mastery, listening, reading comprehension, TOEFL preparation*

INTRODUCTION

In the era of the industrial revolution 4.0, where advancements in technology, science, and information are evolving at an unprecedented pace, the importance of English language proficiency cannot be overstated. English has become a global lingua franca, essential for international communication and the dissemination of scientific knowledge (Crystal, 2003). Mastery of English is not only a prestigious skill but also a fundamental requirement for accessing global academic and professional opportunities.

In ASEAN countries, where English language has status as EFL, including Indonesia, language proficiency tests such as TOEFL serves a key determinant of the readiness of Indonesian university students to engage in academic contexts. Thus, TOEFL results are used as major consideration in the English language skills measurement. TOEFL is intended to assess several academic -related aspects. Firstly, students

are required to have sufficient writing skill and grammar mastery in English in order to be able to write for scientific purpose. Secondly, they are highly demanded to have ability to comprehend English texts and wide range of reading material. Thirdly, students are to have sufficient ability to listen to any descriptions and explanations given in English lectures.

Language proficiency test scores such as TOEFL are claimed as strong predictor of students' academic success in university settings (Dang & Dang, 2023; Ginther & Yan, 2018; Kamaşak et al., 2021; Ruegg et al., 2024; Woodrow, 2014) where English is the language of instruction. Students with higher scores are considered to be better prepared to respond to academic needs. These test scores reflect students' essential language skills (listening, speaking, reading, and writing). Language skills are essentially communication skills so that they can engage effectively such as in understanding coursework, participating in discussions, and completing assignments.

The credibility, validity, and practicality of TOEFL has been long challenged and evaluated in literature, one of which by Dalimunte et al. (2023). They found that TOEFL particularly TOEFL ITP has continued its trustworthiness in the long practices of English proficiency tests. It was claimed that ETS constantly evaluates and improves the credibility, validity, and practicality of their test. A vast body of research has also mainly measured the impacts of proficiency test preparation including TOEFL, IELTS, etc. on learners' English language skills and test outcomes. To begin with, Winke and Lim (2017) investigated the effect of test preparation on L2 listening test performance. They examined the test scores of 169 international students taking the TOEFL listening test after completing an English language program in the United States. Their study found that students who engaged in more test preparation, such as practicing with sample tests and receiving feedback, tended to perform better on the listening test than those who did not prepare as much. They also identified specific types of test preparation activities that were particularly

effective in improving listening skills, such as note-taking and listening to lectures.

Liu (2014) examined the relationship between test preparation and students' performance on TOEFL iBT. Surveying 14,593 respondents in China, data on students' test preparation habits and TOEFL iBT scores were gathered. Liu found that students who engaged in more test preparation, such as taking practice tests, using textbooks, and attending test preparation classes, tended to perform better on the TOEFL iBT than those who did not prepare as much. Mousavi (2014) measured the impact of test preparation on test performance in the IELTS and TOEFL iBT reading tests. Surveying 82 Iranian EFL students taking both reading tests, they showcase that test preparation positively influenced test performance on both tests. Their implication recommended that a combination of both test-taking strategies and language proficiency is critical for test success. In addition, Cohen and Upton (2006, 2007) investigated reading strategies and test-taking strategies used by students when taking the reading section of the new TOEFL test. The observed students' groups consisted of 32 students with diverse first language backgrounds such as Chinese, Japanese, Korean, etc. They found that students used different strategies depending on the format of the questions they faced. Green (2007) compared the impact of IELTS preparation program and pre-university language courses on students' writing ability outcomes. The findings suggested that courses focusing on test preparation tended to be more effective at improving students' test scores compared to pre-university EAP courses that were not intended for the test. Green (2007) also found several individual characteristics that significantly impacted the test outcomes, such as students' confidence, attitude, and motivation when learning English, as well as their previous academic learning experiences.

Studies (i.e., Green, 2007; Liu, 2014; Cohen & Upton, 2007; Mousavi, 2014) indicate that test preparation such as the TOEFL and IELTS plays a crucial role in achieving students' success on English

language assessments. Our review also consistently indicates a positive correlation between test preparation and improved performance on these tests. The reviewed past studies also examined a range of test preparation activities, including practicing with sample tests, utilizing textbooks, taking notes and listening to lectures, and participating in preparation classes. However, one noticeable gap in the literature is regarding the variety of research methods used. Most of the past studies exclusively used quantitative methods to measure any effects of the test preparation. Little study extensively used both quantitative and qualitative methods.

Therefore, to address the methodological gap found in the literature, the present study used mixed method by combining various methods of data collection and analysis both quantitatively and qualitatively. In addition to experiments, our data collection was also carried out through classroom observations. Qualitatively, we used CEFR descriptors to understand in detail the language skills acquired by students as a result of TOEFL teaching interventions. This approach allows us to more accurately analyze the impact of the implemented TOEFL program.

Within Islamic economics and business programs at a university in Indonesia, students are required to achieve a TOEFL score of 475 to meet graduation criteria. However, recent data reveal a concerning trend at the university. Among 757 students taking TOEFL in 2024, only 29% of students scored above 475, while a staggering 71% scored below 475 (Pusat Pengembangan Bahasa UINSU, 2024). This indicates a pervasive issue of inadequate English proficiency among students, which presents substantial barriers to their academic and professional success.

The challenges faced by Indonesian students in achieving satisfactory TOEFL scores are also extensively documented in the literature. Past research explored various challenges that L2 learners encountered when taking to TOEFL test. Literature identified many factors contributing to learners' difficulties such as Lack of structure

and vocabulary mastery (Akmal et al., 2020; Halim & Ardiningtyas, 2018; Kalantar, 2023; Rina & Tiarina, 2021), L1 background (Green, 2007; Kalantar, 2023), Unfamiliar topics (Roohani et al., 2017; Skehan, 2009), motivation (Green, 2007; Kalantar, 2023), lack of teacher's corrective feedback (Ajabshir & Ebadi, 2023; Cheng et al., 2021). These challenges are compounded by socio-demographic factors, such as age and social status, which can influence students' access to resources and opportunities for language practice.

The difficulties students faced in answering TOEFL questions often stem from inadequate foundational English skills. This deficiency is evident in their struggle with grammar, which impacts their ability to tackle questions related to sentence structure and written expression. Past studies showed that participants with low grammar skills struggled to produce complex and accurate sentences (Akmal et al., 2020; Halim & Ardiningtyas, 2018; Kalantar, 2023; Rina & Tiarina, 2021). Participants' tendency to use simple vocabulary and grammatical structures indicated a lack of mastery of the more complex grammar needed to reach the C1 level.

Limited vocabulary mastery also contributes to students' challenges in language test (Akmal et al., 2020; Halim & Ardiningtyas, 2018; Kalantar, 2023). In spite of basic vocabulary mastery in language test, many students have difficulty with it. If many students lack vocabulary mastery, TOEFL instructors can hinder to focus on preparing materials related to more advanced TOEFL strategies and tips. Students' L1 background also plays a role in achieving target levels of proficiency in TOEFL (Green, 2007; Kalantar, 2023). NES students' L1 influence can cause difficulties in using English language accurately and complexly, especially in academic contexts. This condition affects their performance in language proficiency tests.

Previous research also indicated that language learners typically performed better and are able to produce more complex and accurate language when discussing topics with which they are familiar (Roohani et al., 2017; Skehan, 2009). Familiarity with the academic

topics tested on the TOEFL poses a significant challenge for learners. Motivation is strong predictor in a successful language test (Green, 2007; Rina & Tiarina, 2021). Past studies suggested that participants with high motivation tended to be more committed in taking the preparation course and are more engaged in learning and actively practicing to improve their language skills. This has a positive impact on their language test scores. On the other hand, low- motivated participants often viewed the TOEFL test and other language tests as simply a graduation or pass requirement rather than an opportunity to improve their language skills, which in turn results in their low-test scores.

Past research also showed that teacher feedback also plays a significant role in improving the learners' performance (Ajabshir & Ebadi, 2023; Cheng et al., 2021). Corrective feedback from teachers facilitates learners to understand their mistakes and improve their language skills, particularly in grammar, vocabulary, and sentence structure. Lack of sufficient practice also contributes to their difficulties in tackling the test (Kalantar, 2023; Permatasari et al., 2024). Lack of adequate practice in speaking and writing results in low fluency and accuracy. Regular practices, including TOEFL mock tests, are also essential for becoming acquainted with the test's timing and for improving time management across its different sections.

In spite of the presence of a Language Development Center at the aforementioned university, its role has been limited primarily to administering TOEFL tests for the academic community and outsiders. The absence of structured and comprehensive English language proficiency development programs such as TOEFL preparation courses, is a significant oversight. As noted by Green (2007), TOEFL preparation programs are crucial for helping students develop the specific academic language skills necessary for achieving higher scores. Without such support, students are left inadequately prepared, both for the TOEFL and for the broader demands of academic and

professional environments where English is the medium of communication.

Given the mandatory requirement for a TOEFL score of 475 for bachelor's graduation, the current lack of structured language support programs is not only concerning but also indicative of a systemic gap in the university's educational provisions. The implementation of a TOEFL preparation program is thus imperative. Such a program could provide students with targeted instruction in essential language skills, including listening, reading, writing, and speaking, tailored to the specific demands of the TOEFL test (Xie, 2013).

To address the explicated problems, the present study carried out a TOEFL preparation program devoted to Indonesian bachelor's students to prepare them to begin their Islamic economics and business programs. We, thus, hypothesized that the TOEFL program improves their English proficiency particularly at listening skill, structure mastery, and reading skill. Therefore, the present study aimed to measure empirical effect of the TOEFL preparation program on students' English language proficiency at the Islamic economics and business programs in the aforementioned university.

METHOD

The present study used a mixed method approach, combining quantitative and qualitative methodologies to provide a comprehensive analysis of the research problem and findings. The quantitative approach used a quasi-experimental design as the main framework, with data analyzed statistically to measure the effectiveness of the TOEFL intervention. Meanwhile, a qualitative approach was also involved to complement the quantitative findings by providing in-depth analysis of student proficiency levels as result of the test scores. Quasi-experimental research is conducted to determine the effect of manipulation on the observed individual behavior (Allen, 2017; Ary et al., 2014). In educational fields, such as English language learning, experimental research aims to assess the

effect of one specific treatment or action on students' behavior. This type of research also tests hypotheses to find out whether or not one particular action has effect when compared to other actions (Allen, 2017; Ary et al., 2014).

In the present study, one group pretest-posttest design was selected to be applied. The design has two features (Allen, 2017). The first feature is the use of one-group participants i.e., one-group design. This feature indicates that all participants of the group are part of one condition. All participants were given the same intervention and assessment. The second feature is a linear sequence that requires assessment of the dependent variable before and after treatment, namely the pretest-posttest design. In the pretest-posttest research design, the effect of treatment is determined by calculating the difference between the first assessment of the dependent variable through the pretest and the second assessment of the dependent variable through the posttest. Research with one group pretest-posttest design can be illustrated in Figure 1.

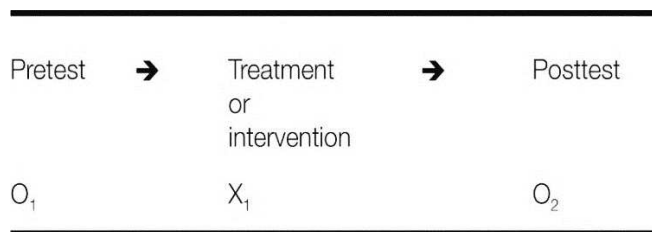


Figure 1. One group pretest-posttest design (Allen (2017)).

One-group pretest-posttest research designs are often used to assess instructors' effectiveness of a given curriculum or course. To achieve this goal, we assessed the students' language knowledge and skills prior to the TOEFL course (i.e., pretest, O_1). Then, the instructors devote their efforts over a period of time to teaching participants and assisting them in acquiring knowledge or skills related to the course topic (i.e., X_1).

At the end of the program, we reassessed the students' knowledge and skills (i.e., posttest, O2). The difference between students' knowledge or skills at the beginning of the course compared to the end of the course is often related to the education provided by the instructor. Therefore, this present investigation was carried out to measure whether or not the implementation of TOEFL preparation affect the English language proficiency of the first-year students of Islamic economics and business programs.

Rationale of the selection of Research Design

Several reasons why one group pre-test post-test was chosen as the design of the present study. Firstly, the design is widely used to evaluate the effectiveness of an educational program. In this study, the effectiveness of the TOEFL preparation program was measured in improving university students' English skills. Secondly, the design is quite simple to be implemented. And the results can be calculated by simple analysis mostly through t-tests. Thirdly, the design is more rigorous than other one-group designs such as the one-group posttest design because it provided a basis for participants' performance. Considering the potential threat of internal validity as described by Phakiti (2014), this present experiment was carried out in a short period of time during seven weeks.

Population and Sample

Population of the study is a source of data based on subjects who deal with the difficulties and are considered to provide complete and accurate information of the investigated problems. The population in this study was the first-year undergraduate students of Islamic economics and business programs. The first-year students were selected because they need to be prepared with sufficient foreign language skills to start their undergraduate studies. Having good language skills such as English reading skill, listening skill, etc. can give them key access to vast sources during their college in order to

learn wide ranges of disciplinary knowledge within their majors such as knowledge on Islamic economic, Sharia banking, accounting, etc. Having background knowledge of the academic content is demanded because TOEFL mostly deal with academic knowledge.

To be member of predetermined population, the students were required to provide some important demographic information on a google form (see the appendix). By filling the form, they showed that they are interested in the TOEFL test and would like to participate voluntarily in the program. Thus, this study involved a total population of 50 students, from which a sample of 20 students was randomly selected to form an intervention group. This group served as the sole participant cohort in the research, with no control group included. Figure 2 shows that their undergraduate majors cover sharia banking (PS), sharia insurance (ASR), sharia accounting (AKS), and sharia economic (EKI).

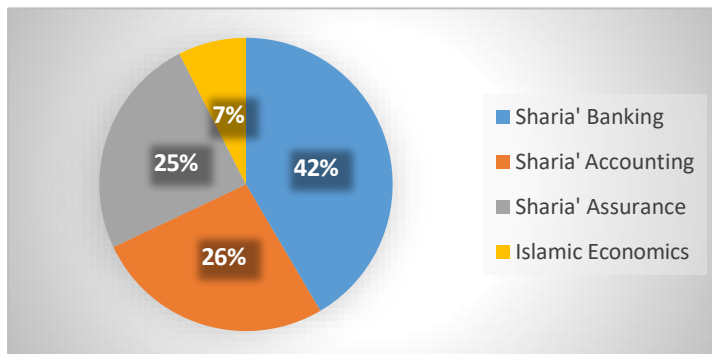


Figure 2. Percentage of undergraduate majors as the population members

This one-group pre-test-post-test research simply involved one group which is an experimental group obtained through randomization technique. Following Federer (1967), sample size of the experimental group was counted to obtain an exact size using formula: $(n-1)(t-1) \geq 15$. The given formula shows that the minimum sample size is 16. Thus, in this present study the sample size of the experimental group was 20. The group consisted of 20 Islamic economics and

business students, with a majority of female participants (15 females and 5 males). The research participants were 2022 cohort between 18 and 19 years old from various Sharia-related fields, with five students each majoring in Sharia Insurance and Sharia Accounting, seven students majoring in Sharia Banking, and three students majoring in Sharia Economics. This group showed a representative sample of students from diverse subfields within Islamic economics and business.

Procedures of Experimental Group Intervention

Following Allen (2017) and Ary et al. (2014), the one-group pretest-posttest design in this present study involved three primary steps. First, the researchers administered a TOEFL pretest measuring the dependent variables of listening, structure and written, and reading performances of the students respectively. Next, the researchers applied experimental treatment to the subjects. The treatment was teaching of TOEFL materials during the program. The TOEFL program was taught by three qualified instructors who are English lecturers holding master's degree in English language education. These instructors possess substantial teaching experience in TOEFL preparation, having taught the subject both at the university level and in language courses. Their academic background and professional expertise enable them to effectively guide students in developing the necessary language skills for the TOEFL exam.

The TOEFL program in this study was dedicated to comprehensive preparation for all sections of the TOEFL test in a short period of time. The program was designed to equip students with English skills and confidence needed to achieve their target scores. The preparation program was conducted over seven weeks outside of their academic hours.

In Week 1, students were introduced to the structure, format, and scoring of the TOEFL test, and a pre-test was administered to diagnose their strengths and areas for improvement. Instruction was

then focused on listening comprehension by teaching strategies for answering questions related to restating factual information, contextual awareness, idiomatic expressions, and logical inference. Limited practice sessions were conducted to improve listening accuracy and speed. In Week 2, listening instruction was continued with a focus on longer conversations and lectures, including identification of main topics, speaker identity, speech purpose, tone, and point of view. Comprehensive practice that included all types of questions was conducted at the end of the week to reinforce understanding and application of the strategies learned. In Week 3, sentence completion and grammar rules in structure and written expression were emphasized. Important grammar concepts, such as subject-verb agreement, tense, and sentence structure, were taught. Timed practice sessions and error analysis were conducted to improve grammatical accuracy. In Week 4, written expression and structure learning was continued with focus on identifying errors in written expression. Students were guided to find incorrect words or phrases in sentences, and advanced grammar rules, including complex sentence structure and adjectives. Comprehensive practice for written expression and structure questions was conducted at the end of the Week 4. In Week 5, reading comprehension skills were developed by teaching strategies for answering questions about ideas, vocabulary, and logical inferences. A variety of question types, including factual information and paraphrasing, were practiced to improve reading speed and comprehension. During Week 6, the focus on reading comprehension was continued to complex passages and critical thinking questions. Timed practice sessions were conducted to improve speed and accuracy, and a full-length practice test was administered to allow students to assess their progress and receive individual feedback.

In the final week, a comprehensive review and reinforcement of all sections of the exam was conducted including test-taking strategies and time management techniques. The program was concluded with a

post-test by researchers to measure the students' progress as the final result of this preparation program. We then evaluated any differences attributed to implementation of the teaching by comparing TOEFL scores of the pretest and posttest.

Data Collection Instruments

The present study used two kinds of instrument of data collection: TOEFL test and classroom observation. The primary instrument of data collection was the administration of TOEFL test. The test was administered two times, before and after the TOEFL treatment. Table 3 reports the TOEFL test sections administered during the pre and posttest based on Philips (2002).

Table 3. Description of TOEFL test sections

Section	Number of Questions	Duration
Listening Comprehension	50 questions	35 minutes
Structure and Written Expression	40 questions	25 minutes
Reading Comprehension	50 questions	55 minutes
Total	140 questions	2 hours

Direct observation was also conducted during each section of the program. Using FIAC (Flanders's Interaction Analysis Categories) classroom observation sheet (Barbadillo et al., 2024), teachers and students' interactions were observed directly. Figure 3 is an example of record sheet of the observation conducted in the present study. The interactions that we observed were divided into ten categories, with categories 1 to 7 representing teacher talk (such as accepting feelings, giving praise, and giving direction), and categories 8 to 10 focusing on learners talk (including learners' response, initiation, and silence).

Data Analysis

The present study carried out two ways of data analysis: statistical data analysis and descriptive analysis using CEFR levels

descriptors. First, the data analysis was conducted through a statistical computer program, SPSS 21. The test given was in a multiple-choice format. The scoring system was based on the TOEFL assessment. Following the commonly applied statistical procedures for experimental research of one group pretest posttest, we conducted several statistical analysis and measurement in our study including descriptive statistics, paired sample t-test and paired sample correlations.

Analysis of TOEFL Score Data based on CEFR Levels

Second, in order to draw more accurate descriptions of the students' English language skill, it is highly important to interpret participants' TOEFL scores based on CEFR or Common European Framework for Languages: Learning, Teaching, and Assessment (Council of Europe, 2001). Table 4 provides benchmark of each level of English proficiency where students with listening comprehension score of less than 38 do not meet the A2 level benchmark proficiency. Those with structure and written expression scores less than 32 do not meet the A2 level benchmark proficiency. Test takers with Reading and Vocabulary score less than 31 have not met the A2 level benchmark proficiency.

Table 4. TOEFL score descriptions based on CEFR levels

CEFR Levels	Proficiency category	Listening Comprehension	Structure and Written Expression	Reading Comprehension	Total Cut Score
C1 Level	Proficient user	64-68	64-68	63-67	627
B2 Level	Independent user	54-63	53-63	56-62	543
B1 Level	Independent user	47-50	43-50	48-50	460
A2 Level	Basic user	38-46	32-42	31-47	337

FINDINGS

The effect of TOEFL preparation program on students' English proficiency using the TOEFL scoring system, the raw score of each TOEFL section was converted to obtain the final score of both the pre-test and post-test. Figure 3 presents the dataset of pre-test and post-test results of the experimental group's participants.

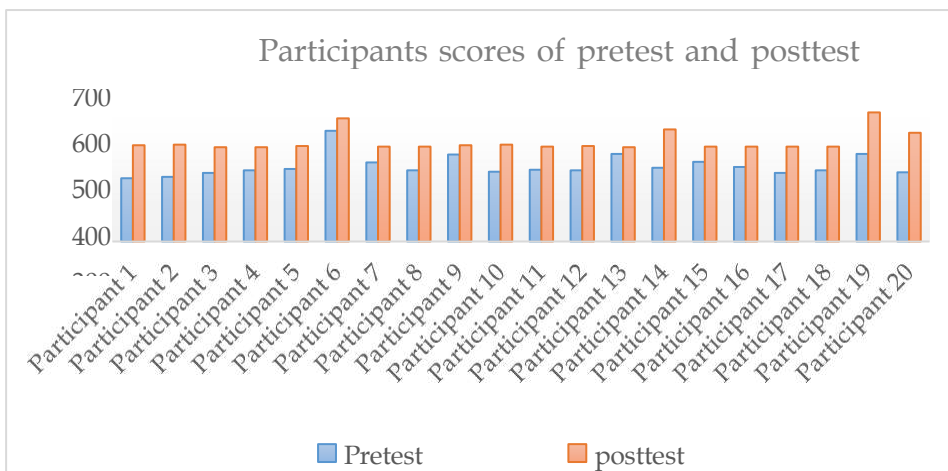


Figure 3. Experimental group's TOEFL score of pre-test and post-test

In the pre-test, participants' scores ranged from 320 to 560. After the intervention, there was an increase in post-test scores ranging from 477 to 653. Each participant experienced an increase in their scores. For example, the first participant who had an initial score of 320 in the pre-test managed to increase his or her score to 487 in the post-test. The participant with the highest score in the post-test was participant number 19, who increased from an initial score of 443 to 653, indicating an increase of 210 points. Overall, these data indicate that TOEFL training has a positive impact on improving participants' English skills.

The results of the statistical measurement show a significant difference in the average score of the experimental group between prior ($M=382.65$, $SD=12.25$) and after the implementation of TOEFL preparation training ($M=505.20$, $SD= 11.52$, $t=11.361$, $p = 0.000$).

Table 5. Paired samples statistics of TOEFL score

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Pre	382,6500	20	54,80038	12,25374
	Post	505,2000	20	51,54926	11,52677

As for t-test calculation in Table 5, it was found that t count is higher than t table ($t \text{ count} = 11.361 > t \text{ table} = 2.09$). Therefore, null hypothesis is rejected and alternative hypothesis is accepted. Although, the standard deviation shows some variations in individual students' scores, the overall results indicate that the TOEFL preparation program had a positive impact on students' English language proficiency.

Table 6. Paired samples correlations of TOEFL score

		N	Correlation	Sig.
Pair 1	Pre & Post	20	,590	,006

Table 6 indicated that correlation value as 0.590. This suggested that participants who performed well on the pretest generally performed well on the posttest. And those who scored lower on the pretest tended to score lower on the posttest. However, the relationship is not perfect. With a p-value of 0.006, our analysis showed strong evidence. It means that the correlation of 0.590 between the pretest and posttest TOEFL scores did not occur by chance. We, therefore, conclude a real relationship between students' performance on the pretest and posttest. This occurred probably due to the effect of the TOEFL preparation program. The probability that this correlation occurred by random chance is only 0.6%.

Results of the paired samples test showed a significant increase in TOEFL scores from pre-test to post-test. The mean score from pre-test to post-test increased with a mean difference of -122.55 (SD = 48.24, SE = 10.79). The statistical test yielded $t(19) = -11.36$ with $p = 0.000$,

indicating that this increase was statistically significant ($p < 0.05$). The 95% confidence interval for the difference ranged from -145.13 to -99.97, confirming that the intervention had a positive impact on improving participants' English language skills in the TOEFL test.

The Effect of TOEFL Program Implementation on Students' Listening skill

Table 8 shows a significant increase in participants' TOEFL listening scores from pre-test ($M = 37.90$, $SD = 5.160$) to post-test ($M = 47.65$, $SD = 5.976$). The modest increased standard deviation indicates variation across participants at post-test. The increased standard error of the mean from 1.154 to 1.336 indicates a modest increase in sample variability in the mean estimates of participants' listening scores between pre-test and post-test.

Table 7. Paired samples statistics of listening score

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Pre	37,90	20	5,160	1,154
	Post	47,65	20	5,976	1,336

Table 8. Paired samples correlations of listening score

		N	Correlation	Sig.
Pair 1	Pre & Post	20	,484	,031

In Table 8, the correlation between pre-test and post-test scores of 0.484 suggested that in general, participants with higher pre-test scores tend to achieve higher post-test scores, although the relationship is moderate. The significance value of 0.031 suggests a significant relationship between participants' listening results on the pre-test and post-test.

The results of the paired samples test for TOEFL Listening scores reported an average increase of 9.75 points from pre-test to post-

test. The standard deviation value of 5.702 indicates an average increase in scores from pre-test to post-test where most participants showed consistent increases in scores. Some participants experienced more significant changes in scores than others. The confidence interval for the difference ranges (from -12.419 to -7.081) indicating a statistically significant increase. The t-value of -7.647 and the significance value ($p=0.000$) indicate a highly significant increase. Thus, these results indicate an increase in participants' listening skills after given the intervention.

To indicate changes in participants' listening comprehension proficiency levels as a result of the preparation program, the following CEFR Figure presents the distribution of test participants' proficiency levels before (pre-test) and after (post-test) the intervention. The data in Figure 4 and Figure 5 are obtained from the results of the pre-test and post-test listening comprehension in the experimental group, respectively. Following the CEFR (Council of Europe, 2001), we categorize the participants' abilities into four CEFR levels, namely A2 (basic user), B1 and B2 (independent), and C1 (Proficient).

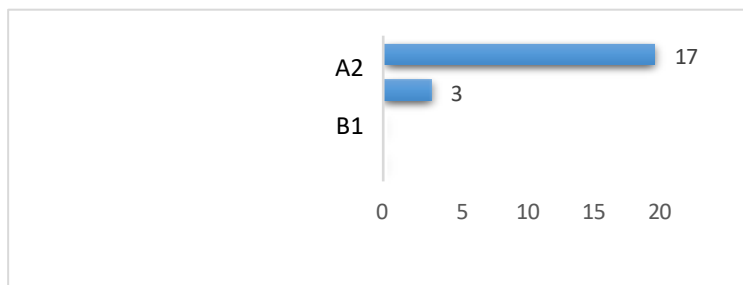


Figure 4. Pre-test result of listening score

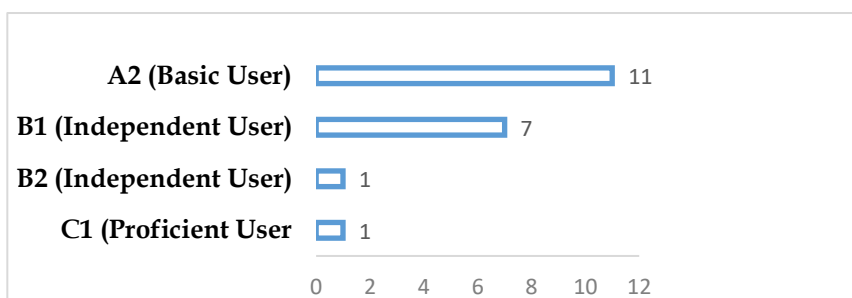


Figure 5. Post-test result of listening score

Figure 5 shows improvements in participants' listening skills after the TOEFL program. Before the intervention, 85% were at A2 level (17 participants) and 15% at B1. Afterward, A2 participants decreased to 55% (11 participants), while B1 increased to 35% (7 participants). Additionally, one participant reached B2 and another C1, levels not observed before. These results indicate progress in listening proficiency based on the CEFR scale.

The Effect of TOEFL Implementation on Students' Structure and Written Expression Mastery

The analysis results showed a significant difference in the students' average scores of structures and written expression mastery between before (M=36.10, SD=1.596) and after the implementation of the TOEFL preparation program (M=51.55, SD=1.725, t=8.662, p = 0.000,).

Table 9. Paired samples correlations of structure and written expression score

		N	Correlation	Sig.
Pair 1	Pre & Post	20	,425	,062

Table 9 specifically shows that t count is higher than t table (t count = 8,662 > t table = 2,09). Thus, H0 is rejected and Ha is accepted. The findings show a difference between the average results of the TOEFL structure pretest and post-test training.

Results of the paired sample test comparing the structure and written scores before and after the intervention yielded the score difference of -15.45 indicating an increase. The standard deviation of 7.977 and the standard error of the mean of 1.784 indicated moderate variability. The confidence interval (-19.183 to -11.717) indicates a statistically significant difference. The t-value of -8.662 and a significance value (p-value) of 0.000 (<0.05) indicates that the increase

in scores on the post-test is statistically significant. Thus, these results indicate an improvement in participants' abilities in structure and written expression after the intervention.

The CEFR results in Figure 6 and Figure 7 respectively illustrates the distribution of test takers' proficiency levels in structure and written expression before (pre-test) and after (post-test) the program. The data were obtained from the participants' TOEFL structure and written expression test results. We found that the participants' structure and written expression abilities were divided into four CEFR levels: A2, B1 and B2, and C1.

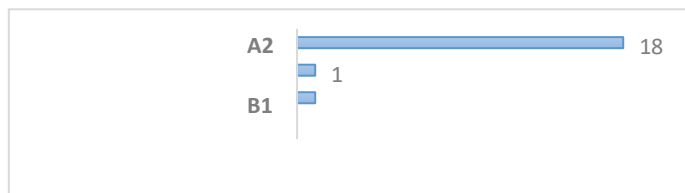


Figure 6. Pre-test result of structure and written expression

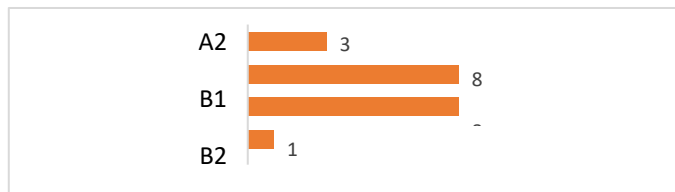


Figure 7. Post-test result of structure and written expression

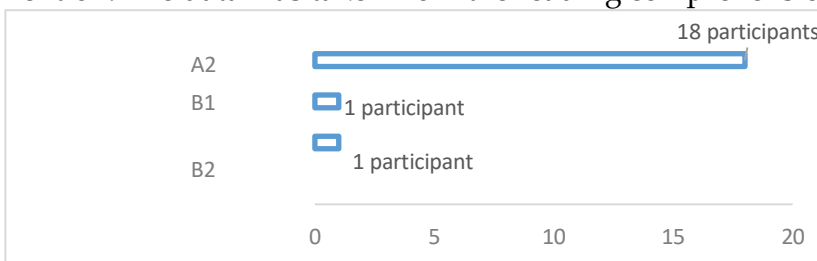
Figure 7 illustrates improvements in structure and written expression skills after the TOEFL program. While 70% of students were at A2 and 30% at B1 before the test, post-test results showed a decrease in A2 to 55%, an increase in B1 to 35%, and the emergence of B2 (5%) and C1 (5%) levels. These results indicate that 45% of students improved their skills, with some reaching higher CEFR levels, demonstrating the program's effectiveness.

The Effect of TOEFL Program Implementation on Students' Reading Skill

The results of statistical measurements showed a significant difference in the mean TOEFL training results (reading) between before ($M=40.80$, $SD= 1.539$) and after implementation ($M=52.35$, $SD= 1.456$, $t= 7.088$, $p = 0.000$). Analysis showed a significant increase in the mean reading scores from pre-test ($M = 40.80$, $SD = 6.89$) to post-test ($M = 52.35$, $SD = 6.52$), with $t(19) = 7.088$, $p = 0.000$. A weak positive correlation ($r = 0.10$, $p = 0.073$) was found between the pre-test and post-test scores. Thus, there is a difference between the mean results of the TOEFL reading pretest and post-test training.

Results also showed a significant increase in the reading scores after the intervention. The average increased in post-test scores was 11.55 points. The standard deviation was 7.28716 indicating variation among participants. The standard error of the mean of 1.62946 illustrates the level of error in estimating the mean difference. The confidence interval (-14.96049 to -8.13951), along with a t-value of -7.088 and a p-value of 0.000 ($p < 0.05$), shows that the increase in scores was statistically significant. This indicates that the given TOEFL preparation successfully improved the participants' reading ability.

The CEFR figures show the distribution of test participants' reading proficiency levels before (pre-test) and after (post-test) the intervention. The data was taken from the reading comprehension test



results and then divided based on their language proficiency levels. The participants' language proficiency was divided into four levels including A2, B1 and B2, and C1.

Figure 8. Pre-test result of reading comprehension proficiency

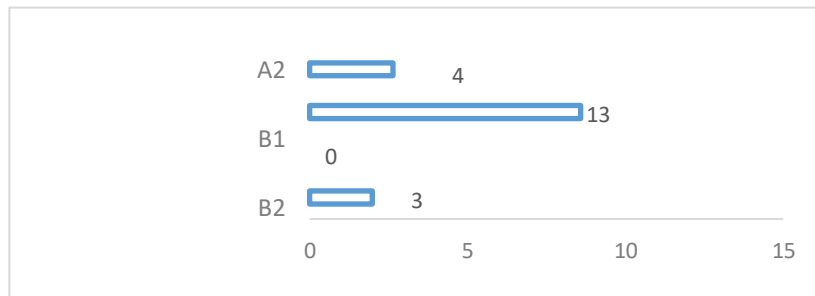


Figure 9. Post-test Result of students' reading comprehension proficiency

Figure 9 shows significant improvements in reading comprehension after the intervention. The percentage of A2-level participants decreased from 90% to 65%, while B1-level participants increased from 5% to 20%. Additionally, B2-level participants rose from 5% to 15%, and 15% of participants reached C1, which was not observed before. These results indicate notable progress, with students advancing to higher reading proficiency levels.

Based on Figure 10, the observed teacher talk was heavily concentrated in categories 4 (asking questions), 5 (giving information), and 6 (giving direction), with counts of 62, 55, and 45 instances respectively. This indicated that the teacher frequently engaged in instructing, providing guidance, and checking students' understanding through questioning. We also recorded fewer instances of accepting ideas or praising (categories 1 and 2). This suggested that the lessons more focused on delivering content than on encouraging student ideas or providing positive reinforcement.

Dalimunte, A. A., Somanawattana, C., & Siregar, D. Y. (2025). The effectiveness of TOEFL program on the students' English proficiency performance: A case of an Indonesian university. *JEELS*, 12(1), 141-173.

Classroom Observation Flanders's Interaction Analysis Categories (FIAC)										
Subject : Reading Instructor's name:						Evaluator: Date : 25 Sept 2022				
CATEGORIES	TEACHER TALK: Response-Initiation							PUPIL TALK: Response-Initiation-silence		
	1	2	3	4	5	6	7	8	9	10
CALCULATION	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00
	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00
	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00
	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00
	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00
	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00
	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00
	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00
	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00
	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00
	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00
	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00
	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00
	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00	00:00
Total Per Categories	64	67	14	62	55	42	45	60	27	45

Notes:	
1. Accepts feeling	6. Giving Direction
2. Praises or encourages	7. Criticizing
3. Accepts or uses ideas of pupil	8. Pupil talk response
4. Asks question	9. Pupil talk initiation
5. Lecturing	10. Silence or confusion

Evaluator's Comments:

Figure 10. Sample sheet of direct classroom observation

On the other hand, the observed learners talk was primarily categorized under response (category 8), with 38 instances. This indicated that students were frequently responding to teacher prompts. However, we also recorded fewer instances of learners' initiation (category 9). This showed that students were not actively leading the interaction. Additionally, 27 instances of silence (category 10) suggested moments of passive participation or pauses. Thus, our classroom observations revealed a teacher-centered interaction activities pattern, with a focus on delivering content and guiding students rather than encouraging a high level of students-led discussion.

DISCUSSION

Our key findings of the present study show that the TOEFL preparation implementation significantly affect Indonesian university students' English language proficiency. The key findings will be discussed in three sub-findings.

The first finding of this study indicates that students' listening comprehension skills increased significantly after participating in the TOEFL preparation program at the Islamic Economics and Business study programs. The results of the data analysis show that the average score of students' listening comprehension ability increased from before the TOEFL preparation program compared to after the program was implemented. Based on our analysis using the CEFR (Council of Europe, 2001), in the average the students are at B1 level where their proficiency is independent user. Students at this level, when listening to a short dialogue, were able to understand high-frequency vocabulary and deduce the meaning of some lower frequency vocabulary. It was also described that students at this proficiency category are considered to understand conversational implications, for example, implied questions given through statements. They also understood various functions of common language, for example, invitations, apologies, suggestions, etc. They also recognized references to various types of pronouns.

Thus, the implication of our findings is that the implementation of the TOEFL preparation program was an effective strategy in improving students' listening comprehension skills at the faculty. The present finding is consistent with Winke and Lim (2017) and Maharani and Putro (2021) who also measured the effect of test preparation on L2 learners' listening ability. Examining foreign students' score of TOEFL listening, they found that students who practiced in the TOEFL preparation program performed better than its counterpart.

The second finding of the present study indicates that students' structure and written expression mastery significantly improves after they were taught TOEFL structure materials. In the average students are at B1 level. Based on the CEFR level descriptors (Council of Europe,

2001), students who achieve this level are usually able to use the common tenses of the verb correctly, including the passive voice of the sentence. Students also recognize when a verb needs an object, such as an infinitive, gerund, or a clause that starts with 'that'. They are also able to identify the correct sentence or clause structure, even if the subject and verb are slightly separated. The finding is consistent with Aprilliandari and Sugiharto (2023) who investigated that the L2 learners improved and perceived written expression after practiced in 10 meetings. However, our finding is not supported by Pranoto (2020) who did not find any improvement on students' structure mastery after the TOEFL training.

The third finding of the present study shows that students' reading comprehension performance significantly improves after they received TOEFL reading materials. Based on the CEFR level, in the average students are in B1 level. In CEFR, it was described that students at this level are usually able to understand descriptions of relatively simple processes and narration in well-marked academic texts. They understood high-frequency vocabulary and recognized paraphrased information. They also followed sentence-level comparisons and contrasts and understand the meaning conveyed by the most common conjunctions. They also connected meaning across some simple sentences that contain high-frequency vocabulary.

The reading strategies were also identified by Cohen and Upton (2007) when comparing TOEFL reading performance in four different NNES students' groups in both East Asia and Middle East. They also listed several reading strategies including returning to the passage, eliminating wrong options, inferencing to understand unfamiliar vocabulary. Our finding is also consistent with Pranoto (2020) who also measured the effectiveness of TOEFL on the enhancement of students' reading comprehension. She found that students' reading skills improved better after the program. Jing and Rahman (2025) conducted a study on effective reading to enhance TOEFL reading scores. Their

findings revealed that learners excelled in academic vocabulary, reading comprehension, and test anxiety management.

Our findings of the improvement of students' English proficiency can also be explained in terms of several aspects including test taking strategy, instructional methods, students' motivation and engagement. Firstly, test-taking strategy is one of key factors influencing the increase in listening and reading comprehension skills. A few past studies such as Permatasari et al. (2024) and Zalha et al. (2020) have emphasized the importance of test-taking strategies to enhance students' test performance. The preparation training taught students several strategies when they encountered listening comprehension tests. First, they attempted to understand the directions of the doing test. Familiarizing with the given listening instructions during the training, they acknowledged saving their time and energy by not reading the instructions carefully as they began the test. Second, students attempted to know where easier and harder questions were usually located. During the TOEFL practices, they were taught that listening questions generally progress from easy to difficult. During the test, they listened carefully to the passages. They concentrated fully on what the speakers were saying on the recording because they can only hear the recording once. They used their remaining time to look ahead to the answers of the following questions. After completed one question, they still had time to look ahead to the answer to the next question.

The next significant aspects to explain how students' performance was facilitated during the TOEFL program is in terms of the teachers' instructions and students' motivation and engagement. Based on classroom observations, in order to keep their students motivated, the teachers were motivated in the first place by showing that they have a strong passion for teaching. Teachers are aware of the central importance of being highly motivated teachers to their student's success in learning English during the TOEFL program. Renandya (2013) emphasized that teachers play a key role to motivate

students. Having high degree of motivation for learning is critical for the Economics and business students. Krashen (1982) claimed that highly motivated learners generally perform better in acquiring second language. Alqurashi (2024) also emphasized the key role of motivation for students' engagement and success in reading activities. He suggested that motivated students are more capable of engaging complex academic texts therefore improving their reading proficiency and overall academic performance. In order to maintain their students' motivation, the TOEFL teachers engaged with the students by involving them within practices and peer discussion. Compassionate, helpful, humorous, and dedicated teachers who genuinely cares about their students' learning and general well-being and establishes strong relationship with them is more likely to be capable of generating an inspiring learning environment and providing constant support to student learning. DÖrnyei (2011) argued that motivated learners are more excited, goal-oriented, dedicated, tenacious, and optimistic in their learning. They are inclined to try to accomplish their target and will not surrender until succeed.

CONCLUSION

The objectives of this research were to measure whether students' English proficiency significantly differed before and after receiving TOEFL instruction. The findings of the present study indicate that TOEFL preparation program significantly affected the students' English language proficiency.

Several limitations in the present study need to be acknowledged. First, the limitation of the study concerns its specific context. This research was conducted only in one specific context; namely, Islamic economics and business students at one university in Indonesia who were preparing for the TOEFL exam. Differences in the research context can affect research results. The effect of the TOEFL preparation program on students' English skills can be different when it is carried out in different learning environments or situations.

Therefore, more extensive research is needed to explore the effects of the TOEFL preparation program in a wider context. The second limitation concerns the sample size of the present study. The study involved a limited sample group, namely one experimental group of Indonesian students in a faculty. Generalization of the findings to a given sample size needs to be exercised with caution. Therefore, further research with a more representative sample size might be able to provide a more comprehensive picture of the effect of the TOEFL preparation program on Indonesian students' English skills.

Several recommendations can be put forward as follows. First, the findings of the present study can be initial findings for the development of a better and more effective TOEFL preparation program in the future. In designing the program, future research needs to pay attention to teaching factors such as quality input, student motivation, developing fluency and using formulaic expressions, as well as providing intensive and continuous instruction. Second, the results of this study can be used as an initial study for curriculum development in the future TOEFL program. The curriculum can be enriched with more diverse and contextual teaching, not only on listening and reading skills but also focusing on developing speaking and writing skills. Third, the findings of this study can also add new insights for further research in a wider context and with more varied methods.

ACKNOWLEDGMENTS

This research was funded by research grants of BOPTN 2022 under the ministry of religious affairs of Indonesia.

DECLARATION OF AI AND AI-ASSISTED TECHNOLOGIES

The authors declare that they have used ChatGPT as AI-assisted technology in the writing process before submission and during revision, but only to improve the language and readability of the

manuscript. The authors take full responsibility for the content, interpretation, and originality of the manuscript.

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), 1–16. <https://doi.org/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>
- Allen, M. (2017). *The SAGE Encyclopedia of Communication Research Methods* (M. Allen, Ed.). SAGE.
- Alqurashi, N. (2024). Exploring how language proficiency and reading anxiety influence reading comprehension in EFL learners: An analytical study of key contributing factors and implications. *Journal of Language and Linguistic Studies*, 20(2), 1–12.
- Aprilliandari, D. I., & Sugiharto, P. A. (2023). TOEFL preparation training program to enhance students' TOEFL score. *Jurnal Education and Development*, 11(1), 487–490. <https://doi.org/10.37081/ed.v11i1.4327>
- Ary, D., Jacobs, L. C., & Sorensen, C. K. (2014). *Introduction to research in education* (9th ed.). Wadsworth, Cengage Learning.
- Barbadillo, D. J. O., Grace Barrete, R. R., Ponce, H. T., & Jay Syting, C. O. (2024). Flanders' interaction analysis category system in action: Analyzing teachers' classroom discourse. *Psych Educ*, 16(6), 637–653. <https://doi.org/10.5281/zenodo.10527875>
- Cheng, X., Zhang, L. J., & Yan, Q. (2021). Exploring teacher written feedback in EFL writing classrooms: Beliefs and practices in interaction. *Language Teaching Research*, 29, 385–415. <https://doi.org/10.1177/13621688211057665>
- Cohen, A. D., & Upton, T. A. (2006). Strategies in responding to the new TOEFL reading tasks. *ETS Research Report Series*, 2006(1), i–162. <https://doi.org/10.1002/j.2333-8504.2006.tb02012.x>
- Cohen, A. D., & Upton, T. A. (2007). "I want to go back to the text": Response strategies on the reading subtest of the new TOEFL®.

Dalimunte, A. A., Somanawattana, C., & Siregar, D. Y. (2025). The effectiveness of TOEFL program on the students' English proficiency performance: A case of an Indonesian university. *JEELS*, 12(1), 141-173.

Language Testing, 24(2), 209–250.
<https://doi.org/10.1177/0265532207076364>

Council of Europe. (2001). Common European Framework of Reference for languages: Learning, teaching, assessment. In *Common European Framework*. Cambridge University Press.
<https://rm.coe.int/1680459f97>

Crystal, D. (2003). *English as a global language* (2nd ed.). Cambridge University Press.

Dalimunte, A. A., Tika, A., Dalimunte, M., Tanjung, A. P., & Suryani, I. (2023). Evaluating TOEFL ITP test: A critical review. *Humanitatis: Journal of Language and Literature*, 9(2), 245–254.
<https://doi.org/10.30812/humanitatis.v9i2.2329>

Dang, C. N., & Dang, T. N. Y. (2023). The predictive validity of the IELTS test and contribution of IELTS preparation courses to international students' subsequent academic study: Insights from Vietnamese international students in the UK. *RELC Journal*, 54(1), 84–98.
<https://doi.org/10.1177/0033688220985533>

Dörnyei, Z. (2011). *Motivational strategies in the language classroom*. Cambridge.

Federer, W. T. (1967). *Experimental theory and application*. Oxford.

Ginther, A., & Yan, X. (2018). Interpreting the relationships between TOEFL iBT scores and GPA: Language proficiency, policy, and profiles. *Language Testing*, 35(2), 271–295.
<https://doi.org/10.1177/0265532217704010>

Green, A. (2007). Washback to learning outcomes: A comparative study of IELTS preparation and university pre-sessional language courses. *Assessment in Education: Principles, Policy and Practice*, 14(1), 75–97. <https://doi.org/10.1080/09695940701272880>

Halim, N., & Ardiningtyas, S. Y. (2018). Difficulties faced by the students in answering TOEFL test questions. *English, Teaching, Learning, and Research Journal*, 4, 219–231.
<https://doi.org/https://doi.org/10.24252/Eternal.V42.2018.A7>

Jing, W., & Rahman, A. B. A. (2025). Effective reading strategies for success in the TOEFL reading section: a comprehensive review. *International Journal of Academic Research in Progressive Education*

Dalimunte, A. A., Somanawattana, C., & Siregar, D. Y. (2025). The effectiveness of TOEFL program on the students' English proficiency performance: A case of an Indonesian university. *JEELS*, 12(1), 141-173.

and *Development*, 14(1), 909-927.
<https://doi.org/10.6007/IJARPED/v14-i1/24622>

Kalantar, O. S. (2023). Attaining C1 level scores in TOEFL iBT: From challenges to needs. *International Journal of Language Testing*, 14(1), 99-113. <https://doi.org/10.22034/IJLT.2023.414966.1287>

Kamaşak, R., Sahan, K., & Rose, H. (2021). Academic language-related challenges at an English-medium university. *Journal of English for Academic Purposes*, 49, 3-41. <https://doi.org/10.1016/j.jeap.2020.100945>

Krashen, S. D. (1982). Principles and practice in second language acquisition. In *TESOL Quarterly*. Pergamon Press. <https://doi.org/10.2307/3586656>

Liu, O. L. (2014). Investigating the relationship between test preparation and TOEFL iBT® performance. *ETS Research Report Series*, 2014(2), 1-13. <https://doi.org/10.1002/ets2.12016>

Maharani, M. S., & Putro, N. H. P. S. (2021). Evaluation of TOEFL preparation course program to improve students' test score. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 25(1). <https://doi.org/10.21831/pep.v25i1.39375>

Permatasari, A. N., Firmansyah, F., & Retnawati, H. (2024). The effect of test preparation on English proficiency performance of English learners: a meta-analysis. *International Journal of Language Testing*, 14(2). <https://doi.org/10.22034/IJLT.2024.432975.1316>

Phakiti, A. (2014). *Experimental research methods in language learning*. Bloomsbury.

Phillips, D. (2002). *TOEFL complete course for the TOEFL test: preparation for the computer and paper tests*. Longman.

Pranoto, Y. H. (2020). The effects of the TOEFL preparation program on reading skills and structure mastery of prospective students. *Journal of English Teaching Adi Buana*, 05(01), 77-88.

Pusat Pengembangan Bahasa UINSU. (2024). *Hasil TOEFL Pusat Pengembangan Bahasa UINSU Medan*. <https://pusatbahasa.uinsu.ac.id>

Renandya, W. A. (2013). Essential factors affecting EFL learning outcomes. *English Teaching*, 68(4), 23-41. <https://doi.org/10.15858/engtea.68.4.201312.23>

Dalimunte, A. A., Somanawattana, C., & Siregar, D. Y. (2025). The effectiveness of TOEFL program on the students' English proficiency performance: A case of an Indonesian university. *JEELS*, 12(1), 141-173.

- Rina, P., & Tiarina, Y. (2021). *An analysis of English department students' difficulties in answering TOEFL test of listening comprehension*. <https://doi.org/10.24036/jelt.v10i1.1113555>
- 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.
- Ruegg, R., Hoang, H., & Petersen, N. (2024). The influence of English language proficiency test scores on the academic success of ESL undergraduate students. *Educational Research and Evaluation*, 29(3-4), 111-129. <https://doi.org/10.1080/13803611.2024.2314533>
- Seyyed Ahmad Mousavi, S. A. N. N. (2014). The effect of test preparation on the test performance the case of the IELTS and TOEFL iBT reading tests. *International Journal of English and Education*, 3(2), 581-584.
- 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>
- Winke, P., & Lim, H. (2017). The effects of test preparation on second-language listening test performance. *Language Assessment Quarterly*, 14(4), 380-397. <https://doi.org/10.1080/15434303.2017.1399396>
- Woodrow, L. (2006). *Academic Success of International Postgraduate Education Students and the Role of English Proficiency*. <https://www.researchgate.net/publication/237441334>
- Xie, Q. (2013). Does test preparation work? implications for score validity. *Language Assessment Quarterly*, 10(2), 196-218. <https://doi.org/10.1080/15434303.2012.721423>
- Zalha, F. B., Alfiatunnur, A., & Kamil, C. A. T. (2020). Strategies in dealing with the reading section of TOEFL prediction: A case of Aceh EFL learners. *IJEE (Indonesian Journal of English Education)*, 7(2), 160-171. <https://doi.org/10.15408/ijee.v7i1.17622>