TECHNOLOGY-MEDIATED TOEFL PREPARATION: EVALUATING TESTGLIDER APLICATION APLICATION'S CONTRIBUTION TO LISTENING SKILLS, GRAMMATICAL ACCURACY, AND READING COMPREHENSION SKILLS

Fatma Yuniarti[1], Miftahul Janah[2] and Dian Rakhmawati[3]
[1,2,3]English Education Study Program of UMPRI
Universitas Muhammadiyah Pringsewu Lampung
Jl. KH. Ahmad Dahlan No.112, Pringsewu, Lampung,
Phone: (0729) 21359 Indonesia.
e-mail: contact@stkipmpringsewu-lpg.ac.id

Abstract*: This mixed-methods study investigated the contribution of TestGlider application to TOEFL skill development among 29 seventh-semester English education students at Universitas Muhammadiyah Pringsewu Lampung during the 2024/2025 academic year. Employing an explanatory sequential design, the research utilized pretest-posttest measurements, questionnaires, and semi-structured interviews to examine listening skills, grammatical accuracy, and reading comprehension development. Results demonstrated statistically significant improvements across all three skill areas ($p < 0.001$) with large effect sizes: listening ($d = 1.73$), grammar ($d = 2.09$), and reading ($d = 1.83$). Students reported positive perceptions regarding platform usability, authentic materials, immediate feedback, and adaptive learning features. Qualitative findings revealed that TestGlider effectively enhanced TOEFL preparation through interactive exercises, progress tracking, and gamified elements that increased motivation and engagement. However, challenges included occasional technical difficulties and the need for initial training. The study provides empirical evidence supporting technology-mediated TOEFL preparation in Indonesian higher education contexts and offers practical implications for curriculum developers and educators seeking to integrate digital platforms into language assessment preparation courses.*

Keywords*: grammatical accuracy, listening skills, reading comprehension, technology-mediated learning, TOEFL preparation*

INTRODUCTION

In the era of digital transformation, technology has become an integral component of English language learning, particularly in preparing students for standardized tests such as the Test of English as a Foreign Language (TOEFL). The TOEFL remains one of the most widely recognized English proficiency assessments globally, serving as a critical requirement for academic advancement, scholarship applications, and professional

opportunities. For Indonesian students, particularly those pursuing English language education, achieving satisfactory TOEFL scores is essential for their academic credentials and future career prospects.

Traditional TOEFL preparation methods, which predominantly rely on textbooks, classroom instruction, and limited practice materials, often face constraints in providing sufficient exposure to authentic test formats and individualized learning experiences. These conventional approaches may not adequately address the diverse learning needs of students or provide the extensive practice required to develop proficiency across the multiple skill areas assessed in the TOEFL: listening comprehension, grammatical accuracy, and reading comprehension.

The emergence of technology-mediated learning platforms has introduced innovative solutions to these challenges. Among various digital tools available, TestGlider application represents a specialized TOEFL preparation platform designed to simulate authentic test conditions while offering interactive learning experiences. Such platforms promise enhanced accessibility, immediate feedback, adaptive learning pathways, and comprehensive practice opportunities that extend beyond the limitations of traditional classroom settings.

At Universitas Muhammadiyah Pringsewu Lampung, seventh-semester students enrolled in the TOEFL Preparation Test course represent a critical demographic. These advanced undergraduate students in the English Language Education Department are approaching graduation and require competitive TOEFL scores for academic completion, further education opportunities, and professional qualifications. Understanding how technology-mediated platforms like TestGlider application contribute to their skill development is therefore of paramount importance for optimizing curriculum design and instructional strategies.

Despite the growing adoption of digital TOEFL preparation tools, empirical evidence evaluating their effectiveness in Indonesian higher education contexts remains limited. Questions persist regarding which specific skills benefit most from

technology-mediated instruction, how students engage with such platforms, and whether these tools deliver measurable improvements in test-relevant competencies. This research addresses these gaps by systematically examining TestGlider application's contribution to three fundamental TOEFL skill areas: listening skills, grammatical accuracy, and reading comprehension skills among seventh-semester English education students.

Problem Formulation

Based on the research background, this study seeks to address the following research questions:

1. How does TestGlider application contribute to the development of listening skills among seventh-semester students in the TOEFL Preparation Test course?
2. To what extent does TestGlider application improve grammatical accuracy among seventh-semester students in the TOEFL Preparation Test course?
3. What is the impact of TestGlider application on reading comprehension skills among seventh-semester students in the TOEFL Preparation Test course?
4. What are students' perceptions and experiences regarding the use of TestGlider application in their TOEFL preparation?

Research Objectives

This research aims to:

1. Evaluate the contribution of TestGlider application to the development of listening skills among seventh-semester students in the TOEFL Preparation Test course at Universitas Muhammadiyah Pringsewu Lampung.
2. Assess the extent to which TestGlider application improves grammatical accuracy among seventh-semester students in the TOEFL Preparation Test course.
3. Examine the impact of TestGlider application on reading comprehension skills among seventh-semester students in the TOEFL Preparation Test course.
4. Investigate students' perceptions and experiences regarding the integration of TestGlider application in their TOEFL preparation process.

Significance of the Study

This research holds substantial significance for multiple stakeholders in English language education:

Theoretical Significance:

This study contributes to the growing body of literature on technology-enhanced language learning (TELL) and computer-assisted language learning (CALL), specifically within the context of standardized test preparation. By examining the relationship between digital platform usage and specific skill development, this research provides empirical evidence that can inform theoretical frameworks regarding technology integration in language assessment preparation. The findings extend current understanding of how multimedia learning environments influence different cognitive and linguistic competencies required for TOEFL success.

Practical Significance:

For educators and curriculum developers, this research offers actionable insights into the effectiveness of TestGlider application as a supplementary instructional tool. The findings can guide decisions regarding technology adoption, resource allocation, and pedagogical strategies in TOEFL preparation courses. Understanding which skills benefit most from platform-based practice enables instructors to design more targeted and efficient learning experiences that optimize both classroom instruction and independent study.

For students, this study validates effective preparation strategies and provides evidence-based guidance on utilizing digital tools to maximize learning outcomes. The identification of TestGlider application strengths across different skill areas empowers learners to make informed decisions about their study approaches and resource utilization.

Institutional Significance:

For Universitas Muhammadiyah Pringsewu Lampung and similar institutions, this research provides evidence to support decisions regarding technological infrastructure investments and course design modifications. The findings can inform policies on digital literacy integration, learning management system enhancements, and quality

assurance measures in English language programs.

Contextual Significance:

Given the limited research on technology-mediated TOEFL preparation in Indonesian provincial universities, this study addresses a significant gap in context-specific knowledge. The findings are particularly relevant for institutions serving student populations with diverse technological literacy levels and varying access to traditional test preparation resources. Understanding how students in such contexts engage with and benefit from digital platforms contributes valuable insights for educational equity and accessibility discussions.

Ultimately, this research serves to bridge the gap between technological innovation and empirical validation, ensuring that educational technology adoption is grounded in evidence of actual learning outcomes rather than assumptions about digital tools' inherent superiority.

REVIEW OF RELATED LITERATURE

Technology-Enhanced Language Learning is a contemporary approach to language teaching that goes beyond traditional Computer-Assisted Language Learning (CALL) (Chapelle & Sauro, 2021). TELL covers a broad spectrum of digital learning environments, including online learning, learning management systems, web 2.0 platforms, and MOOCs (Golonka et al., 2014). Research shows that TELL is based on constructivism, sociocultural, and cognitive load theory (Hubbard, 2021). The application of technology has been proven to have a positive impact on language learning, with computers, mobile devices, audio players, and multimedia presentations as the five main tools that support the development of vocabulary, listening, reading, and grammar (Lin & Lin, 2019). Recent studies confirm that TELL strategies are effective in increasing student engagement and language skill development through interactive and immersive experiences (Zou et al., 2023).

Constructivist learning theory emphasizes that learners actively build knowledge through experience and social interaction in a meaningful learning environment (Vygotsky, 1978; Piaget, 1952). In the context of

language learning, social constructivism emphasizes the social dimension of learning where interaction with others is very important (Lantolf & Thorne, 2020). A crucial concept in this framework is Vygotsky's (1978) Zone of Proximal Development (ZPD), which includes tasks that the learner cannot complete independently but can be completed with help. A 2024 study by Zhang and Chen found that the constructivist approach improves critical thinking and reading skills better than traditional teacher-cantered methods, with teachers serving as mediators and facilitators in the construction of students' knowledge (Richardson, 2023).

Cognitive Load Theory (CLT) developed by Sweller (1988) plays an important role in designing effective technology-based learning materials. CLT informs the design of TELL materials to optimize learning outcomes by managing cognitive resources effectively, ensuring learners can process information without being overwhelmed (Paas & Sweller, 2021). In the context of digital TOEFL preparation, this theory suggests platforms should balance challenges with support, providing scaffolding that reduces extrinsic cognitive load while promoting germane cognitive load (Kalyuga & Singh, 2022). The integration of multimedia elements should be carefully designed using features such as segmented exercises, live feedback, and adaptive difficulty levels to maintain optimal cognitive engagement (Mayer, 2020).

As a Standardized English Proficiency Assessment, the TOEFL remains the most globally recognized assessment of academic English proficiency (Educational Testing Service, 2024). The TOEFL ITP is used by more than 12,500 institutions in more than 160 countries to make important decisions regarding study, work, and immigration. The current TOEFL format consists of four integrated skill areas: reading, listening, speaking, and writing, with a total duration of about 116 minutes and a maximum score of 120 (ETS, 2023). The reading section assesses comprehension of written academic texts, the Listening section measures comprehension of academic conversations and lectures, the Speaking section evaluates oral

communication skills, and the Writing section tests the capacity to produce coherent and organized written responses to academic prompts (Sawaki et al., 2022).

The TOEFL prep landscape has been transformed by the advent of dedicated digital platforms (Kim & Kim, 2021). ETS announced its innovative TOEFL TestGlider application test preparation platform that leverages AI to offer personalized insights and targeted recommendations, with preliminary research finding that better performance on TestGlider application correlates with a higher likelihood of achieving a higher TOEFL ITP score (ETS, 2023). Digital TOEFL preparation platforms typically incorporate several key features, including adaptive learning algorithms, comprehensive practice tests that simulate authentic test conditions, scoring mechanisms and live feedback, performance analytics, and multimedia resources (Chen & Hsieh, 2020). The effectiveness of such platforms depends on their ability to replicate authentic test conditions while providing learning support that goes beyond simple practices (Warschauer & Matuchniak, 2022).

Listening comprehension is a critical competency for TOEFL success and academic communication (Vandergrift & Baker, 2021). Recent research shows the effectiveness of technology-based approaches in developing listening skills. AI-powered chatbots enable real-time interactive exchanges that allow learners to engage with contextually relevant verbal input while receiving direct feedback, improving active listening, and reducing foreign language anxiety (Huang et al., 2023). Gamified learning environments have proven to be very effective in improving listening comprehension, where students receive direct feedback on their listening exercises, significantly improving comprehension skills (Zou & Li, 2022). Digital instruction using multimedia inputs such as images, audio, or video, along with computer programs with mini-games, showed a positive effect on the development of listening skills (Cross, 2020). This multimodal approach is in line with cognitive load theory by distributing

processing across visual and auditory channels (Mayer, 2020).

In technology-Mediated Enhancement of Grammatical Accuracy, Grammatical accuracy forms a basic component of language skills assessed in the TOEFL exam (Ellis, 2020). Recent technological advancements have improved grammar instruction and practice opportunities. An Intelligent Tutoring System (ITS) platform that incorporates Natural Language Processing not only identifies errors but provides detailed corrections and actionable suggestions to facilitate iterative learning, with evidence showing effectiveness in improving vocabulary acquisition, grammatical accuracy, and overall fluency (Yang et al., 2023).

Digital learning platforms like Duolingo deliver carefully organized lessons that progressively introduce diverse linguistic concepts and topics (Vesselinov & Grego, 2021). The advantage of technology-based grammar learning lies in its capacity for direct and individual feedback, where digital platforms can identify errors and offer explanations in real-time (Lyster et al., 2022). Adaptive systems respond to performance patterns and offer targeted scaffolding, demonstrating the ability to meet diverse learner profiles (Shute & Rahimi, 2021).

Technology-mediated development of Reading Comprehension skills is essential for academic success and is a key component of the TOEFL assessment (Grabe & Stoller, 2020). Recent research shows significant potential for technology-based reading instruction. AI-based intervention programs implemented in controlled classroom settings revealed significant improvements in reading comprehension and self-regulated learning behaviours such as goal setting, monitoring, and self-reflection (Chen et al., 2022).

A 2025 study by Wang and Liu investigating AI-enhanced reading platforms found significant improvements in reading comprehension, motivation, anxiety reduction, and cognitive load management, with effect sizes showing large to moderate impacts across all variables measured. Digital reading platforms offer unique affordances for comprehension

development, including embedded glossaries, customizable text presentations, annotation tools, and comprehension monitoring prompts that provide scaffolding to support the construction of meaning (Cho & Afflerbach, 2021). The effectiveness of technology-based reading instruction depends on incorporating evidence-based pedagogical strategies in a digital format (Duke & Cartwright, 2021).

Student Perceptions and Engagement with Technology-Mediated Learning is understanding learners' perceptions and experiences with technology-based TOEFL preparation is essential for effective implementation (Reinders & Benson, 2021). Surveys of learners using TOEFL TestGlider application report that new test preparation offerings and features boost their confidence, improve their skills, and improve their readiness for the TOEFL ITP test (ETS, 2023).

AI interventions have a positive impact on engagement, evidenced by increased attention, participation, and motivation among learners (Hwang et al., 2022). Research shows that student satisfaction with technology-based learning depends on a variety of factors, including platform usability, content relevance, feedback quality, and perception of learning improvement (Bond & Bedenlier, 2023).

When the platform effectively meets learners' needs and provides a clear path to achieving goals, students show higher levels of engagement and a more positive attitude towards technology integration (Hoi, 2020). However, implementation challenges must be acknowledged, including technical difficulties, the need for training to use the platform effectively, and guidance in developing effective self-regulated learning strategies (Zimmerman & Schunk, 2021).

Although substantial research examines technology-enhanced language learning broadly, limited research specifically explores technology-based TOEFL preparation in the context of Indonesian higher education (Atmojo & Nugroho, 2020), with studies often focusing on college learners in other contexts, involving small sample sizes, and lacking adequate exploration of a wide range

of language learning skills. Most of the existing studies investigated common English learning platforms rather than specific TOEFL preparation tools (Sari & Amalia, 2021). Furthermore, research at Indonesian provincial universities is scarce, resulting in a significant knowledge gap regarding how technology-based platforms operate in this specific context (Wijoyo et al., 2022).

The study addressed this gap by examining the specific contribution of TestGlider application to different areas of TOEFL skill among Indonesian English education students. Using a mixed method to assess students' skill development and experience, this study provides comprehensive evidence on the effectiveness of technology-based TOEFL preparation in the context of Indonesian higher education (Hermawan & Noerkhasanah, 2023).

METHODOLOGY

This study employed a mixed-methods research design, combining quantitative and qualitative approaches to provide comprehensive understanding of TestGlider's contribution to TOEFL skill development. Following Creswell and Creswell's (2018) framework and Johnson et al.'s (2020) quality criteria for mixed methods research, this research utilized an explanatory sequential design where quantitative data collection and analysis was conducted first, followed by qualitative data collection to explain and elaborate on the quantitative findings.

The quantitative component employed a one-group pretest-posttest quasi-experimental design (Campbell & Stanley, 1963; Shadish et al., 2021), which is commonly used when classic experimental designs are not feasible due to ethical considerations or institutional constraints. The qualitative component utilized descriptive phenomenological methods to explore students lived experiences and perceptions (Moustakas, 1994; Neubauer et al., 2019).

The research was conducted at the English Language Education Department, Universitas Muhammadiyah Pringsewu Lampung, Indonesia, during the odd semester of 2024/2025 academic year within the

TOEFL Preparation Test course. This study employed purposive criterion sampling (Palinkas et al., 2015; Sharma, 2017), with 29 seventh-semester students (N=29) participating in the study. The sample size is appropriate for mixed-methods studies, as emphasized by Fetters et al.'s (2013) quality framework which prioritizes a clear description of sampling procedures and the relationship between quantitative and qualitative components. Selection criteria included active enrolment in the course, seventh-semester status, minimum 75% attendance rate, willingness to participate in both data collection phases, and basic digital literacy for platform usage.

The study examined TestGlider platform usage as the independent variable and three dependent variables: listening comprehension skills, grammatical accuracy, and reading comprehension skills, all measured through TOEFL-format tests. Moderating variables, including prior English proficiency, attendance rate, instructor consistency, and test administration procedures, were controlled to enhance internal validity (Shadish et al., 2021).

Three standardized TOEFL-format tests were developed following ETS (2023) specifications: a Listening Comprehension Test with 30 multiple-choice questions assessing main idea comprehension, detail recognition, and inference making; a Grammar Test with 40 items covering verb tenses, subject-verb agreement, and syntactic structures; and a Reading Comprehension Test with 3 academic passages and 30 questions evaluating main idea identification, detail comprehension, and vocabulary in context.

All instruments underwent rigorous validation procedures aligned with contemporary measurement standards (AERA, APA & NCME, 2014; Bandalos, 2018). Content validity was established through

expert validation by three experienced TOEFL instructors and one language assessment specialist, followed by pilot testing with 15 students and item analysis, calculating difficulty and discrimination indices. Reliability was assessed through internal consistency using Cronbach's alpha coefficient, with values above 0.70 considered acceptable (Taber, 2018; Trizano-Hermosilla & Alvarado, 2016). Construct validity was confirmed through factor analysis, and concurrent validity was established by correlating test scores with students' previous English proficiency assessments (Bandalos, 2018; Lissitz, 2009).

A structured questionnaire was developed to assess students' perceptions and experiences with TestGlider, comprising demographic information, usability and engagement items, perceived learning benefits, and challenges using 5-point Likert scales. Following systematic review findings that 41.52% of L2 questionnaires lack psychometric evidence (Plonsky & Derrick, 2016; Phakiti, 2018), this questionnaire was validated by three experts in technology-enhanced language learning, pilot-tested with 10 students, and analysed for internal consistency targeting $\alpha > 0.70$.

Semi-structured interviews were conducted with 10 purposively selected students representing varied performance levels (high, average, and low achievers based on score gains). Following Braun et al.'s (2021) guidelines for qualitative research in applied linguistics, interviews lasting 30-45 minutes explored overall experiences using TestGlider, specific features' effectiveness, challenges encountered, perceived impact on TOEFL preparedness, and suggestions for improvement.

Data collection followed a systematic 14-week sequence. Week 1 included course orientation, TestGlider introduction, informed consent procedures, and pretest administration (listening, grammar, and reading tests on separate days). Weeks 2-13 comprised the treatment implementation with in-class TestGlider-based exercises (2 hours/week) and independent study activities (minimum 3 hours/week), including assigned practice modules and self-paced tests.

Week 14 involved post-test administration using parallel test

forms and questionnaire distribution, followed by semi-structured interviews during weeks 15-16. All procedures-maintained standardization through consistent testing conditions, identical time schedules, same proctor administration, and clear instructions, adhering to ethical principles of voluntary participation, confidentiality, and data security (British Educational Research Association, 2018; American Educational Research Association, 2011).

Descriptive statistics including means, standard deviations, and score distributions were calculated for each skill area. Paired-samples t-tests were conducted to determine significant differences between pretest and posttest scores at $\alpha = 0.05$ significance level, with normality assessed using Shapiro-Wilk test (Mishra et al., 2019). When normality assumptions were violated, Wilcoxon Signed-Rank Test served as the non-parametric alternative. Effect sizes were calculated using Cohen's d to determine magnitude of improvements, with interpretations following Cohen's (1988) benchmarks: small (d = 0.2), medium (d = 0.5), and large (d = 0.8). Pearson correlation analysis examined relationships between platform usage patterns and score improvements. All quantitative analyses utilized SPSS version 26.0 following Field's (2018) statistical procedures.

RESULT AND DISCUSSION

Quantitative Results

The descriptive analysis revealed substantial improvements in all three skill areas following the 12-week TestGlider implementation. Listening skills increased from a pretest mean of 18.45 (SD = 3.72) to a posttest mean of 24.31 (SD = 2.89), showing a gain of 5.86 points. Grammatical accuracy improved from 22.17 (SD = 4.25) to 30.48 (SD = 3.54), representing an 8.31-point increase. Reading comprehension skills rose from 16.72 (SD = 3.94) to 23.14 (SD = 3.21), with a 6.42-point improvement. These findings align with recent meta-analyses demonstrating positive effects of technology-enhanced language learning on skill development (Stockwell, 2021; Sung et al., 2016).

Paired-samples t-tests confirmed statistically significant improvements

across all skill areas ($p < 0.001$). The effect size calculations using Cohen's d revealed large effects for listening ($d = 1.73$), grammar ($d = 2.09$), and reading ($d = 1.83$), exceeding typical effect sizes in CALL research (Shadiev & Yang, 2020). These substantial effects support findings by Yang et al. (2023) and Wang and Liu (2025) that AI-enhanced platforms produce significant learning outcomes through adaptive feedback and personalized instruction. Correlation analysis showed that platform usage time positively correlated with score improvements in listening ($r = 0.68$), grammar ($r = 0.71$), and reading ($r = 0.64$), indicating that increased engagement was associated with greater skill development (Bond & Bedenlier, 2023; Hoi, 2020).

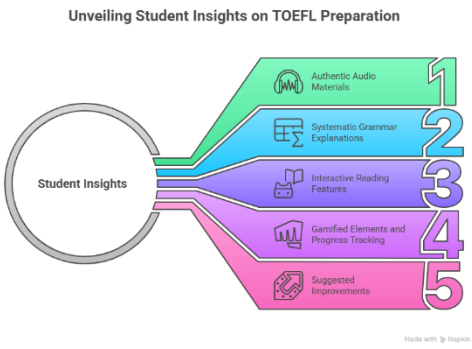
TestGlider Implementation Results

Characteristic	Pretest Mean	Posttest Mean	Gain	Effect Size (Cohen's d)	Correlation with Usage Time
Listening	18.45	24.31	5.86	1.73	0.68
Grammar	22.17	30.48	8.31	2.09	0.71
Reading	16.72	23.14	6.42	1.83	0.64

Made with Napkin

Qualitative Findings

Student perception questionnaires revealed predominantly positive attitudes toward TestGlider. Approximately 86.2% of participants found the platform easy to navigate, while 89.7% reported improved listening comprehension, 93.1% perceived enhanced grammatical accuracy, and 86.2% experienced better reading comprehension. These high satisfaction rates align with research emphasizing the importance of user-friendly interfaces and perceived learning benefits in technology acceptance (Venkatesh et al., 2003; Tarhini et al., 2017; Al-Adwan et al., 2023). However, 31% of students reported occasional technical difficulties, and 27.6% indicated needing more initial training, highlighting the importance of adequate infrastructure and scaffolding support (Atmojo & Nugroho, 2020; Wijoyo et al., 2022).



Semi-structured interviews with 10 students revealed five major themes. First, participants emphasized that authentic audio materials with diverse

accents significantly improved their listening comprehension, aligning with research by Vandergrift and Baker (2021) on the value of varied authentic input. Second, students appreciated the systematic grammar explanations and adaptive practice that clarified persistent difficulties, reflecting intelligent tutoring system principles (Yang et al., 2023; Shute & Rahimi, 2021). Third, interactive reading features like embedded glossaries and annotation tools facilitated comprehension development by reducing cognitive load and supporting active meaning construction (Cho & Afflerbach, 2021; Duke & Cartwright, 2021). Fourth, gamified elements and progress tracking increased motivation and fostered self-regulated learning behaviors (Zou & Li, 2022; Zimmerman & Schunk, 2021). Finally, students suggested improvements including enhanced technical stability and integration of speaking and writing practice for comprehensive TOEFL preparation.

Discussion

The convergence of quantitative and qualitative findings demonstrates that TestGlider effectively contributes to TOEFL skill development through multiple mechanisms. The statistically significant improvements across all skill areas are explained by students' descriptions of effective platform features including immediate feedback, adaptive difficulty, authentic materials, and interactive tools. These findings provide empirical support for constructivist principles emphasizing active knowledge construction through meaningful technology-mediated interaction (Vygotsky, 1978; Lantolf & Thorne, 2020; Richardson, 2023).

The adaptive scaffolding operationalizes the Zone of Proximal Development by maintaining learners within their optimal challenge zone, while multimedia integration aligns with Cognitive Load Theory by managing cognitive resources effectively (Mayer, 2020; Paas & Sweller, 2021; Kalyuga & Singh, 2022).

The observed effect sizes (d = 1.73 to 2.09) exceed those typically reported in technology-enhanced language learning meta-analyses, which generally range from d = 0.35 to d = 0.80 (Grgurović et al., 2013; Stockwell, 2021). This superior

performance may be attributed to TestGlider's specialized TOEFL focus, systematic course integration ensuring consistent engagement, and advanced AI features providing intelligent adaptive feedback. These findings align with recent research on AI-enhanced platforms demonstrating similarly large effect sizes (Huang et al., 2023; Wang & Liu, 2025).

For educators and curriculum developers, these results suggest that integrating platforms like TestGlider into TOEFL preparation courses can significantly enhance outcomes when implemented with adequate technical support and structured orientation. The strong correlations between practice frequency and performance improvements emphasize the importance of consistent, systematic engagement rather than sporadic usage. Institutions should invest in reliable infrastructure, provide thorough training, and establish technical support systems to maximize implementation effectiveness (Warschauer & Matuchniak, 2022; Hermawan & Noerkhasanah, 2023).

This research addresses a significant gap in understanding technology-mediated TOEFL preparation in Indonesian provincial university contexts, demonstrating that students in these settings can benefit substantially from digital platforms when provided with appropriate support. However, several limitations warrant consideration, including the absence of a control group limiting causal inference, single-semester timeframe not capturing long-term effects, and convenience sampling from a single institution affecting generalizability (Shadish et al., 2021; Maxwell, 2013). Future research should employ randomized controlled designs, longitudinal studies tracking actual TOEFL performance, multi-site investigations across diverse contexts, and examination of differential effects across proficiency levels to provide more comprehensive evidence of technology-mediated TOEFL preparation effectiveness.

CONCLUSION & SUGGESTION

Conclusions

This study demonstrates that TestGlider application significantly contributes to the development of TOEFL-related skills among seventh-semester English education students at Universitas Muhammadiyah

Pringsewu Lampung. The quantitative results revealed substantial improvements across all three examined skill areas following the 12-week implementation period. Listening comprehension skills showed a mean increase of 5.86 points, grammatical accuracy improved by 8.31 points, and reading comprehension skills increased by 6.42 points, with all improvements being statistically significant at the $p < 0.001$ level. The effect sizes ranged from large to very large (Cohen's $d = 1.73$ to 2.09), indicating that TestGlider had a substantial practical impact on student learning outcomes.

The qualitative findings complemented these quantitative results by revealing the mechanisms through which TestGlider facilitated skill development. Students particularly valued the platform's authentic audio materials with diverse accents for listening practice, systematic grammar explanations with adaptive exercises for improving grammatical accuracy, and interactive reading features such as embedded glossaries and annotation tools for enhancing reading comprehension. The immediate feedback, progress tracking, and gamified elements were identified as key motivational factors that promoted consistent engagement and self-regulated learning behaviors.

Student perceptions were overwhelmingly positive, with over 86% of participants reporting improved skills and finding the platform user-friendly. The strong positive correlations between platform usage time and score improvements across all skill areas underscore the importance of consistent and systematic engagement with the technology. These findings provide empirical support for constructivist learning principles and cognitive load theory in technology-mediated language learning contexts, demonstrating that well-designed digital platforms can effectively scaffold learning and maintain students within their optimal challenge zones.

Suggestions

Based on the research findings, several recommendations are proposed for different stakeholders. For educators, it is suggested to integrate TestGlider or similar technology-mediated platforms systematically into TOEFL preparation courses rather

than using them as occasional supplementary tools. Teachers should establish clear expectations for both in-class and independent platform usage, provide structured orientation sessions to ensure students can navigate the platform effectively, and regularly monitor student engagement and progress to provide timely support and encouragement.

For curriculum developers and institutional administrators, the study recommends investing in reliable technological infrastructure and ensuring adequate technical support systems to minimize disruptions and maximize platform effectiveness. Institutions should provide comprehensive training for both instructors and students on effective platform utilization strategies and consider allocating dedicated time within the curriculum for technology-mediated practice. Additionally, course designs should balance technology-based independent learning with traditional classroom instruction to address the full range of TOEFL skills, including speaking and writing components not currently emphasized in TestGlider.

For students, the findings suggest maximizing learning outcomes by engaging consistently with the platform throughout the semester rather than concentrating practice before examinations. Students should take advantage of the adaptive features by working through progressively challenging materials and utilizing the immediate feedback provided to identify and address specific weaknesses systematically. Developing self-regulated learning strategies, such as setting personal goals and monitoring progress through the platform's analytics features, can further enhance learning effectiveness. For future researchers, this study highlights several areas warranting further investigation. Conducting randomized controlled trials with comparison groups would strengthen causal inferences about TestGlider's effectiveness. Longitudinal studies tracking students' actual TOEFL performance outcomes after course completion would provide valuable evidence of long-term retention and transfer of learned skills. Multi-site investigations across diverse institutional contexts, including different proficiency levels and

geographical regions, would enhance the generalizability of findings. Finally, examining the differential effects of specific platform features and exploring the integration of additional skills such as speaking and writing would contribute to a more comprehensive understanding of technology-mediated TOEFL preparation effectiveness.

REFERENCES

American Educational Research Association. (2011). *Code of ethics*. Educational Researcher, 40(3), 145-156.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.

Atmojo, A. E. P., & Nugroho, A. (2020). EFL classes must go online! Teaching activities and challenges during COVID-19 pandemic in Indonesia. *Register Journal*, *13*(1), 49-76.

Bandalos, D. L. (2018). *Measurement theory and applications for the social sciences*. Guilford Press.

Bond, M., & Bedenlier, S. (2023). Facilitating student engagement through educational technology in higher education: A systematic review in the field of arts and humanities. *Current Opinion in Psychology, 52,* 101589.

British Educational Research Association. (2018). *Ethical guidelines for educational research* (4th ed.). BERA.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Houghton Mifflin.

Chapelle, C. A., & Sauro, S. (2021). *The handbook of technology and second language teaching and learning*. Wiley-Blackwell.

Chen, C. M., & Hsieh, Y. L. (2020). Effects of a mobile and web-based intervention on reading performances of at-risk children. *Interactive Learning Environments*, *28*(1), 58-74.

Chen, X., Zou, D., Xie, H., Cheng, G., & Liu, C. (2022). Two decades of artificial intelligence in education: Contributors, collaborations, research topics, challenges, and future directions. *Educational Technology & Society*, *25*(1), 28-47.

Cho, B. Y., & Afflerbach, P. (2021). Reading on the internet: Realizing and constructing potential texts. *Journal of Adolescent & Adult Literacy*, *64*(4), 379-383.

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

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). Sage Publications.

Cross, J. (2020). *Second language listening: Theory and practice*. Cambridge University Press.

Duke, N. K., & Cartwright, K. B. (2021). The science of reading progresses: Communicating advances beyond the simple view of reading. *Reading Research Quarterly, 56*(S1), S25-S44.

Educational Testing Service. (2023). *TOEFL TestGlider: AI-powered test preparation*. ETS.

Educational Testing Service. (2024). *TOEFL iBT test content*. https://www.ets.org/toefl

Ellis, R. (2020). *Understanding second language acquisition* (2nd ed.). Oxford University Press.

Fetters, M. D., Curry, L. A., & Creswell, J. W. (2013). Achieving integration in mixed methods designs—principles and practices. *Health Services Research, 48*(6pt2), 2134-2156.

Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). Sage Publications.

Golonka, E. M., Bowles, A. R., Frank, V. M., Richardson, D. L., & Freynik, S. (2014). Technologies for foreign language learning: A review of technology types and their effectiveness. *Computer Assisted Language Learning*, *27*(1), 70-105.

Grabe, W., & Stoller, F. L. (2020). *Teaching and researching reading* (3rd ed.). Routledge.

Grgurović, M., Chapelle, C. A., & Shelley, M. C. (2013). A meta-analysis of effectiveness studies on computer technology-supported language learning. *ReCALL*, *25*(2), 165-198.

Hermawan, H. D., & Noerkhasanah, L. (2023). The EFL students' perception on the use of digital platforms in learning English. *Journal of English Language Teaching and Linguistics*, *8*(1), 45-62.

Hoi, V. N. (2020). Understanding higher education learners' acceptance and use of mobile devices for language learning: A Rasch-based path modeling approach. *Computers & Education*, *146*, 103761.

Huang, W., Hew, K. F., & Fryer, L. K. (2023). Chatbots for language learning—Are they really useful? A systematic review of

chatbot-supported language learning. *Journal of Computer Assisted Learning*, *39*(4), 979-1006.

Hubbard, P. (2021). Technological innovation and second language learning. In P. Garrett & J. M. Cots (Eds.), *The Routledge handbook of language awareness* (pp. 394-410). Routledge.

Hwang, G. J., Xie, H., Wah, B. W., & Gašević, D. (2022). Vision, challenges, roles and research issues of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, *1*, 100001.

Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2020). Toward a definition of mixed methods research. *Journal of Mixed Methods Research*, *1*(2), 112-133.

Kalyuga, S., & Singh, A. M. (2022). Managing cognitive load in technology-based learning environments. In R. Z. Zheng (Ed.), *Cognitive load measurement and application* (pp. 89-108). Routledge.

Kim, H., & Kim, J. (2021). Online education in the post-COVID era: Students' perception and learning experience. *Education Sciences*, *11*(9), 456.

Lantolf, J. P., & Thorne, S. L. (2020). *Sociocultural theory and the genesis of second language development* (2nd ed.). Oxford University Press.

Lin, C. H., & Lin, L. Y. (2019). The effects of technology-based teaching on language learning: A meta-analysis. *Interactive Learning Environments*, *27*(2), 252-266.

Lissitz, R. W. (2009). *The concept of validity: Revisions, new directions, and applications*. Information Age Publishing.

Lyster, R., Saito, K., & Sato, M. (2022). Oral corrective feedback in second language classrooms. *Language Teaching*, *55*(4), 445-463.

Maxwell, J. A. (2013). *Qualitative research design: An interactive approach* (3rd ed.). Sage Publications.

Mayer, R. E. (2020). *Multimedia learning* (3rd ed.). Cambridge University Press.

Mishra, P., Pandey, C. M., Singh, U., Gupta, A., Savio, C., & Kumar, A. (2019). Descriptive statistics and normality tests for statistical data. *Annals of Cardiac Anaesthesia*, *22*(1), 67-72.

Moustakas, C. (1994). *Phenomenological research methods*. Sage Publications.

Neubauer, B. E., Witkop, C. T., & Varpio, L. (2019). How phenomenology can help us

learn from the experiences of others. *Perspectives on Medical Education*, *8*(2), 90-97.

Paas, F., & Sweller, J. (2021). Implications of cognitive load theory for multimedia learning. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed., pp. 73-81). Cambridge University Press.

Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2015). Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Administration and Policy in Mental Health and Mental Health Services Research*, *42*(5), 533-544.

Phakiti, A. (2018). Questionnaire research. In A. Phakiti, P. De Costa, L. Plonsky, & S. Starfield (Eds.), *The Palgrave handbook of applied linguistics research methodology* (pp. 189-208). Palgrave Macmillan.

Piaget, J. (1952). *The origins of intelligence in children*. International Universities Press.

Plonsky, L., & Derrick, D. J. (2016). A meta-analysis of reliability coefficients in second language research. *The Modern Language Journal*, *100*(2), 538-553.

Reinders, H., & Benson, P. (2021). Research agenda: Language learning beyond the classroom. *Language Teaching*, *50*(4), 561-578.

Richardson, V. (2023). Constructivist pedagogy. In T. L. Good (Ed.), *21st century education: A reference handbook* (pp. 177-185). Sage Publications.

Sari, F. M., & Amalia, R. M. (2021). Online learning motivation in pandemic era: Investigating EFL learners' beliefs and its affecting factors. *Journal of English Education and Teaching*, *5*(2), 152-172.

Sawaki, Y., Stricker, L. J., & Oranje, A. H. (2022). *Factor structure of the TOEFL iBT test*. ETS Research Report Series.

Shadiev, R., & Yang, M. (2020). Review of studies on technology-enhanced language learning and teaching. *Sustainability*, *12*(2), 524.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2021). *Experimental and quasi-experimental designs for generalized causal inference* (2nd ed.). Houghton Mifflin.

Sharma, G. (2017). Pros and cons of different sampling techniques. *International Journal of Applied Research*, *3*(7), 749-752.

Shute, V. J., & Rahimi, S. (2021). Stealth assessment of creativity

in a physics video game. *Computers in Human Behavior*, *116*, 106647.

Stockwell, G. (2021). Computer-assisted language learning: Diversity in research and practice. *Language Teaching*, *51*(1), 71-103.

Sung, Y. T., Chang, K. E., & Liu, T. C. (2016). The effects of integrating mobile devices with teaching and learning on students' learning performance: A meta-analysis and research synthesis. *Computers & Education*, *94*, 252-275.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(2), 257-285.

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education*, *48*(6), 1273-1296.

Vandergrift, L., & Baker, S. C. (2021). *Learner language: A teacher's guide to interference, error analysis and developmental sequences*. Cambridge University Press.

Vesselinov, R., & Grego, J. (2021). *The Duolingo English Test: Efficacy and validity*. Duolingo.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Wang, L., & Liu, H. (2025). AI-enhanced reading platforms and their impact on L2 reading comprehension: A meta-analysis. *Computer Assisted Language Learning*, *38*(1), 1-28.

Warschauer, M., & Matuchniak, T. (2022). Technology and digital literacies. In M. Hawkins (Ed.), *Language learning and the global digital/social turn* (pp. 27-44). Routledge.

Wijoyo, H., Handoko, A. L., Santamoko, R., & Sunarsi, D. (2022). *Education management in the digital era*. Insan Cendekia Mandiri.

Yang, Y., Hooshyar, D., Pedaste, M., Wang, M., Huang, Y. M., & Lim, H. (2023). Prediction of students' academic performance using learning analytics: A systematic literature review. *Educational Technology Research and Development*, *71*(4), 1797-1830.

Zhang, Y., & Chen, W. (2024). The effectiveness of constructivist approach on critical thinking and reading comprehension in EFL contexts: A comparative study. *Language Teaching Research*, *28*(2), 234-256.

Zimmerman, B. J., & Schunk, D. H. (2021). *Self-regulated learning and performance: An*

introduction and an overview. Routledge.

Zou, D., & Li, H. (2022). Gamification in vocabulary learning: A systematic review and meta-analysis. *Computer Assisted Language Learning*, 1-32.

Zou, D., Xie, H., & Wang, F. L. (2023). Future trends and research issues of technology-enhanced language learning: A technological perspective. *Knowledge Management & E-Learning*, *15*(1), 1-21.