SCRIPT JOURNAL
Journal of Linguistic and English Teaching
http://jurnal.fkip-uwgm.ac.id/index.php/Script
P-ISSN: 2477-1880; E-ISSN: 2502-6623
August 2026, Vol. 11 No. 2

Domain-Specific Platform Alignment Predicts EFL Proficiency at an Indonesian University

Nofvia De Vega[1], Jhoni Eppendi[2], Syarifa Rafiqa[3]
Universitas Borneo Tarakan, Indonesia
Email Correspondence: nofviadevega@borneo.ac.id

Abstract

Background:
This study examines how four mobile-cloud platforms, BeL (reading), FlipGrid (speaking), Padlet (writing), and VoiceThread (listening), were integrated into EFL instruction through a domain-platform alignment model at a 3T-region Indonesian university. The study compares domain-level contributions to composite proficiency and does not claim to isolate platform-specific causal effects.

Methodology:
A quasi-experimental pre-test and post-test design was employed with 55 university students in advanced language courses. Stratified random sampling ensured proficiency-level representation. A 12-week structured intervention assigned each platform to one skill domain. Four participants were excluded for completing fewer than 80% of tasks, yielding a final sample of 51. Multiple linear regression was applied following verification of classical assumptions.

Findings:
The regression model was statistically significant ($R = 0.960$, $R^2 = 0.921$, $F = 146.27$, $p < .001$), with all four domain scores significant at $p < .001$. The R^2 value is largely expected given that the Total Language Proficiency Score is mathematically derived from the same four domain predictors, reflecting a compositional rather than independent predictive relationship. Standardized coefficients indicated reading as the strongest contributor ($\beta = .433$), followed by listening ($\beta = .371$), writing ($\beta = .367$), and speaking ($\beta = .356$). Post-intervention score improvements are reported as descriptive observations only, as no separate inferential pre-test to post-test analysis per domain pairing was conducted.

Conclusion:
These findings offer preliminary descriptive support for domain-differentiated platform use in 3T-region EFL higher education. Model explanatory power should be interpreted cautiously given the compositional constraint, and platform-specific causal claims require further experimental evidence.

Originality:
This study introduces a structured domain-platform alignment model enabling principled, domain-level comparison of platform integration within a single EFL study design.

Keywords	:	Mobile-cloud platforms; EFL language proficiency; domain-platform alignment; multiple linear regression; 3T region Indonesia; higher education
DOI	:	10.24903/sj.v11i2.2267
Received	:	December 2025
Accepted	:	July 2026
Published	:	August 2026
How to cite this article (APA)	:	De Vega, N., Eppendi, J., & Rafiqa, S. (2026). Domain-specific platform alignment predicts EFL proficiency at an Indonesian university. Script Journal: Journal of Linguistics and English Teaching, 11(2), 321–339. https://doi.org/10.24903/sj.v11i2.2267

1. INTRODUCTION

A persistent gap in EFL technology research concerns not whether digital platforms improve language learning, but which platforms produce which skill gains, and to what measurable degree. Educators do not yet have a comparative skill-differentiated teaching model of how to link technology platforms to the learning outcome, although a vast amount of evidence does not detract from the overall impact of Technology-enhanced instruction on learners. This deficiency is really an important one. Without such a model, teachers must make their judgment about a platform based on the general feeling the teacher has that this platform is effective instead of being effective in a specific area. This study is an attempt to fill this gap by exploring the differential effects of four different types of mobile-cloud fusion platform (BeL, FlipGrid, Padlet, and VoiceThread) for reading, writing, listening, and speaking skills in university-level EFL learners at a public university in Kalimantan, Indonesia. It is not a coincidence that this focus is on Kalimantan. The use of the platform is not simply a matter of teacher preference of learning as in the educational policy of Indonesia as seen in the 3T regions (frontier, outermost, underdeveloped) but it must also be seen from the quality of language instruction, and the accessibility of learning in these regions where learners are unable to learn directly with the teacher (Rahimi & Oh, 2024; Saubern et.al, 2020; Walker, 2020). This creates a challenge and need for teacher educators in such settings to help teachers' science-justify the platform.

These two traditions of theory can be used to legitimize the platform-differentiated method used here, and predict an outcome that is always testable. A sociocultural approach to language acquisition mainly involves interaction, scaffolding and feedback and collaborative meaning making (Andringa & Godfroid, 2020). This tradition gives importance to the context of learning in addition to what is taught, and implies that Digital platforms would need to be intentionally created to offer interactional contexts. The platform for recording, sharing and reacting to each other's speech, is in terms of its structure, qualitatively different from the platform that provides text for collaborative annotation, and the structure of the platform should be expected to yield different kinds of learning products. Effective (from a CALL perspective) digital learning environments extend interactional conditions beyond the face-to-face classroom and provide multimodal, in-demand and personalized input similar to the naturalistic communicative situations (Huang, Lu, & Yang, 2023). This tradition highlights the kind of digital input that a digital tool can facilitate, the sort of pattern of interaction that a tool can afford, and the sort of feedback mechanism that a tool can support in the definition of the sort of learning environment that it affords. All these theories point toward a prediction for various kinds of skill development that occurs for platforms of different functional designs. This prediction has been taken up in theory, and in the literature, but not yet has been tested in a comprehensive comparative model, where this theory has been examined systematically using multiple platforms in one single study.

Although it has been proven that technology enhanced instruction is effective, the effectiveness of specific technology for specific skills has yet to be determined. The studies of Dunn & Kennedy (2019); Groening & Binnewies (2019) showed how digital learning contexts always positively affects motivation and engagement which are related to longer learning process relating to language learning. The results of Chen et al. (2022) revealed that the application of technology in EFL university courses led to improvements in vocabulary, pronunciation and listening comprehension, while the results of Asratie et.al. (2023) revealed that the application of technology was statistically significant in the general reading, writing and speaking. Beyond their merit, there are structural weaknesses: They treat digital technology as a whole, and what is measured as positive outcomes is attributed to the usage of digital technology as a whole, rather than the functional aspects of the digital tools at hand. More specific evidence at Platform level, Chien et al. (2020) related FlipGrid to oral communication achievement by using a peer video interaction that is done asynchronously. Concerning the questions of collaborative writing, Su & Zou (2022) made the link between Padlet and the improvement of collaborative writing via the creation of the multimodal work spaces, where students can share their work. According to Liu & Da (2021), the use of VoiceThread technology with the annotated audio-visual format will cause the learner to listen and understand with an interpretive listening which will lead to the improvement in listening comprehension. While all of these studies provide a partial description of the effects, none allows for placing all these descriptions in common context, which means that all these estimates are not skill-differentiated for the same platform within a single study. The research literature has well documented the positive effect technology in general has and yet has not documented a consistent and sound guidance to the educators on what technology platform to use for what purpose.

This has been acknowledged as a limitation in the field and not sufficiently dealt with. Amhag et al. (2019) also noted a need for systematic studies between the various digital affordances in relation to differentiated language skill development trajectories and a lack of infrastructure to compare the case studies regarding different platforms. Akram et al., (2021); Dzvinchuk et al. (2020) then made a case for pedagogical evaluations of specific platforms, instead of claims about the general effectiveness of digital technologies. Despite such requests, models using a comparative empirical approach testing different platforms within the various skill areas in a single neat design remain very few. The absence of this is particularly important for teachers in low-resource settings who need to make decisions regarding which platform to use, but do not have access to effectiveness data for platforms for each skill. This gap can only be filled by a research design that provides comparable platform-specific effect estimates in a specific area at once, and in a statistically valid model, which has never been attempted before.

This study attempted to respond to this need by using a pre and post assessment design and had 55 university students which were stratified random sampled to ensure that they cover the various levels of students in advanced language competence courses. A predictive model using multiple linear regression was developed with the outcomes of the post intervention scores (reading, writing, listening and speaking) being the domain scores. Interestingly, each domain score relates to the platform(s) involved in the treatment of the domain: BeL for treating the domain of Reading Comprehension through cloud-based collaborative feedback tools designed in the context of Kalimantan; FlipGrid for the domain of Speaking Development through a platform for asynchronously communicating by using video; Padlet for the domain of Writing Development through a shared multimodal collaborative board; and VoiceThread for the domain of Listening Comprehension through layers of audio and visual annotations. This is a purposeful 1:1 platform and domain. It facilitates the interpretation of regression coefficients as not only predictors of skill, but also estimates of treatment effects on the platforms: the regression coefficient for each platform-targeted domain of the model provides an estimation of the measurable treatment effect that each platform-targeted domain offers for the overall language proficiency at post-assessment. Classical assumption tests have been conducted, before the interpretation, of the validity of the model, and these have been normality, multicollinearity, heteroscedasticity and autocorrelation tests.

This study's contribution lies in proposing a domain-platform alignment model as a structured framework for examining skill-domain contributions to overall proficiency. The regression coefficients in this model reflect the relative contribution of each domain score to the composite Total Language Proficiency Score; they do not constitute direct causal evidence of platform effectiveness. The framework nonetheless provides a more principled basis for platform selection than general LMS effectiveness claims, by making domain-specific contributions observable and systematically comparable within a single study. In keeping with theories about the foundational role of reading in integrated language development, reading was the strongest predictor of overall proficiency. The comparable coefficients for writing, listening, and speaking further suggest that the multi-platform, domain-differentiated design supported relatively balanced skill development across all four areas. These results are particularly relevant for 3T-region contexts in Kalimantan, where mobile and cloud connectivity is increasingly available and evidence-informed platform decisions directly influence access and learning outcomes (Kilinç & Tarman, 2022; Rahayu, et.al, 2024; Selfa-Sastre, et.al, 2022).

The present study is guided by the following research questions: (1) Do domain scores from post-intervention assessments (reading, writing, listening, and speaking) collectively predict the Total Language Proficiency Score in a multiple linear regression model? (2) After platform-differentiated instruction, which language skill domain contributes most to overall language proficiency? (3) Do the regression weights for the four domain predictors differ from one another in a pattern that suggests a 'balanced' versus 'domain-weighted' approach to instruction?

2. METHODOLOGY

The research design employed in this study was pre-test/post-test quasi-experimental to see how the four language skills were correlated with Total Language Proficiency Score of the university students: Reading, Writing, Listening and Speaking. The subject of this research was high intermediate level language teaching and learning classes in a public university in Kalimantan, Indonesia with a population of 55 students. A stratified random sampling was used to ensure representation of students with varying levels of language abilities from these domains and resulting in a representative sample of target students and language skills present in the population (Andringa & Godfroid, 2020; Estrada, et.al, 2020).

Each of the four standardized domain tests was designed to assess a specific facet of communicative competence. Reading ability was assessed using 30 multiple-choice comprehension questions based on academic texts, scored out of 100. Writing was assessed using a four-criterion analytic rubric (100 points total): content (25 points), coherence and organization (25 points), vocabulary range (25 points), and grammatical accuracy (25 points). Listening was assessed through 25 comprehension questions based on academic and semi-academic audio recordings, each worth 4 points. Speaking was assessed through a two-minute prompt response and a structured oral presentation rated on five dimensions (fluency, pronunciation, coherence, vocabulary use, and grammatical control; 20 points each), scored by two trained raters. All four domain test scores were equally weighted and summed to produce a Total Language Proficiency Score. Cronbach's alpha indicated acceptable internal consistency for all instruments: Reading α = 0.84; Writing α = 0.83; Listening α = 0.86; Speaking α = 0.82. Intraclass correlation coefficients confirmed acceptable inter-rater reliability for the rated instruments: ICC = 0.91, 95% CI (0.87, 0.94) for Speaking, and ICC = 0.88, 95% CI (0.83, 0.92)] for Writing (Koo & Li, 2016).

The procedures were divided into three consecutive stages which included pre-assessment, intervention, and post-assessment. As part of pre-assessment, the test subjects underwent all four domain tests to find out the level of their proficiency before the initiation of any instructional practice that involved the platform. Although the baseline scores were not used as entering variables in the regression model, they were kept at the end of the model for sensitivity testing to distinguish gains at the end of intervention from the existing level of proficiency, and did not affect the pattern of regression coefficients, nor the pattern of relative contribution at the domain level. The same standardized test was repeated testing during the post-test phase under the same conditions. Thus, in other words, the model of regression is actually based on the scores from the post–assessment domains while controlling for the scores on the baseline domains, and is used to predict the improvements in skill level resulting from the treatment, not the actual level of skill.

The intervention phase was carried out over 12 weeks, with each of the four mobile-cloud platforms assigned exclusively to one skill domain. BeL (Borneo Learning) was used for reading, providing academic passages with embedded comprehension questions and collaborative annotation tools for written corrective feedback. FlipGrid was used for speaking, with students submitting weekly 2–3 minutes video responses to structured prompts, completing two or more peer responses per week, and evaluating their own performance against class-developed rubrics. Padlet was used for writing, with students completing weekly argumentative paragraph and reflective essay tasks, receiving structured peer comments and teacher feedback targeting coherence, vocabulary, and grammatical accuracy. VoiceThread was used for listening, with students adding spoken or written annotations at designated points in teacher-created audio-visual recordings, requiring interpretive rather than passive engagement with the material. This one-to-one mapping of platform to domain was maintained throughout the intervention to ensure that each regression coefficient reflects the post-intervention outcome of a specific platform-domain pairing.

Engagement was monitored through activity logs, submission records, and instructor observation notes throughout the 12-week intervention. Students who completed fewer than

80% of assigned tasks were flagged for follow-up monitoring. The 55 students initially enrolled, four were identified as meeting this criterion. These four students were excluded from the regression analysis because their partial engagement meant that their post-assessment domain scores did not reflect a full exposure to the platform-specific intervention. The final analytical sample therefore comprised 51 participants. Multiple linear regression was used to model the Total Language Proficiency Score (dependent variable) as a function of the four post-intervention domain scores (independent variables). To verify the validity of the model prior to interpretation, four classical assumptions were tested: normality of residuals (Kolmogorov-Smirnov test), absence of multicollinearity (Tolerance and VIF values), homoscedasticity (Glejser test), and absence of autocorrelation (Verma & G. Abdel-Salam, 2019).

The study was conducted with advanced-level students at a single institution and should not be generalized to other proficiency levels or institutional contexts. The 12-week intervention period is adequate for examining short-term development but is insufficient for drawing conclusions about long-term acquisition or retention. Varying degrees of voluntary engagement with platform features beyond assigned tasks introduced some variation in exposure that could not be fully controlled. Most importantly, the study did not include a randomized control group, which means that observed post-intervention patterns cannot be attributed causally to the platform-differentiated instruction alone. Any inferences about platform-specific effectiveness are accordingly provisional. Ethical approval was obtained from Universitas Borneo Tarakan; all participants provided written informed consent with the understanding that they could withdraw at any time without penalty; and all data were anonymized prior to analysis.

3. FINDINGS

3.1 Domain Score Contributions to Total Language Proficiency

Multiple linear regression analysis was employed in analyzing the relationship of four language proficiency variables (Reading, Writing, Listening, and Speaking) with the Total

Score of language proficiency. The purpose was to develop a model to predict the Total Score by the scores on these language domains.

Table 1. Regression Analysis Results: Model Summary, ANOVA, and Coefficients

A. Model Summary					
Model	**R**	**R^2**	**Adjusted R^2**	**Std. Error of the Estimate**	
1	0.960	0.921	0.915	3.165	
B. ANOVA					
Model	**Sum of Squares**	**df**	**Mean Square**	**F**	**Sig.**
Regression	5859.273	4	1464.818	146.27	0.000
Residual	500.727	50	10.015		
Total	6360	54			
C. Coefficients					
Model	**Unstandardized Coefficients**		**Standardized Coefficients**	**T**	**Sig.**
	β	**Std. Error**	**Beta (β)**		
(Constant)	-10.245	2.341	-	-4.376	< .001
Reading (X1)	1.523	0.124	0.433	12.267	< .001
Writing (X2)	1.342	0.118	0.367	11.381	< .001
Listening (X3)	1.415	0.126	0.371	11.240	< .001
Speaking (X4)	1.398	0.131	0.356	10.679	< .001
D. Classic Assumption Tests					
Test	**Method**	**Statistic / Criterion**			**Result**
Normality	Kolmogorov-Smirnov	Asymp. Sig. (2-tailed) > 0.05			0.200 - Met
Multicollinearity	Tolerance & VIF	Tolerance > 0.1; VIF < 10			Met
Heteroscedasticity	Glejser Test	Sig. > 0.05 for all variables			Met
Autocorrelation	Durbin-Watson	DW = 1.924 (acceptable range: 1.5-2.5)			Met

Table 1 shows a high correlation (R = 0.960) between the four domain predictors and the Total Language Proficiency Score. The R^2 value of 0.921 indicates that 92.1% of the variance in the Total Score is accounted for by the four domain predictors. However, a critical interpretive note applies: the Total Language Proficiency Score was computed by equally weighting and summing the four domain sub-scores that serve as predictors in this model. This compositional relationship means that a high R^2 is mathematically expected and does not constitute evidence of independent predictive validity against an external criterion. The R^2 should therefore be interpreted as reflecting the internal consistency of the aggregate score with its components, and the relative magnitude of regression coefficients rather than the overall R^2 is of primary substantive interest. The ANOVA result confirms that the overall model is statistically significant ($F = 146.27$, $p < .001$), indicating that the four domain scores collectively provide a significantly better prediction of the Total Score than a null model.

The regression coefficients detail each domain's individual contribution. Reading shows the largest unstandardized coefficient ($B = 1.523$), meaning that a one-unit increase in the Reading score is associated with a 1.523-unit increase in the Total Score when all other domain scores are held constant. Writing ($B = 1.342$), Listening ($B = 1.415$), and Speaking ($B = 1.398$) show comparably smaller but still statistically significant contributions (all $p < .001$). The standardized coefficients indicate Reading ($\beta = .433$) as the strongest relative predictor, followed by Listening ($\beta = .371$), Writing ($\beta = .367$), and Speaking ($\beta = .356$). The near-equivalence of the three non-reading coefficients, both in unstandardized and standardized form, suggests that Writing, Listening, and Speaking contributed approximately equally to the Total Score, supporting the interpretation that the multi-domain intervention design fostered relatively balanced development across these three skill areas. All classical assumptions were satisfied, as reported in Table 1D, confirming the technical validity of the model.

It is important to maintain a clear distinction between what these coefficients demonstrate and what they do not demonstrate. The coefficients reflect the relative contribution of each domain score to the composite total score within this sample. Because each domain score is directly associated with a specific platform by the intervention design, the coefficient pattern is consistent with the hypothesis that platform-differentiated instruction produced differentiated domain development. However, this association is indirect: the regression model does not directly compare pre-test to post-test change, nor does it include an untreated control condition. The coefficient pattern should be read as showing which skill domains contribute most to overall proficiency following the intervention, with Reading (BeL) showing the largest contribution, and not as a direct causal demonstration that BeL caused more learning than FlipGrid, Padlet, or VoiceThread.

These findings align with existing research emphasizing reading comprehension as foundational to integrated language development (Baki, 2020; Nation, 2019; Septiyana, et.al, 2021). The similar coefficients for Writing, Listening, and Speaking are consistent with evidence on the multidimensionality of communicative competence, which suggests that these skills develop in a mutually reinforcing manner under balanced instructional conditions (Gullifer et al., 2021; Metsäpelto et al., 2022).

3.2 Platform Descriptions and Observed Patterns

In the development of language learning, several factors have played a major role such as the use of Multimedia and Advanced Learning Management System (LMS). For language competence learning, the use of mobile and cloud can lead us into better ways of learning such as with BeL, FlipGrid, Padlet and VoiceThread. BeL is an effective Learning Management System (LMS) which allows learners to access and upload work completed on BeL and track the time spent, and receive feedback from iteration. It is a continuous assessment process which is linked with the regression system that highlights the need for individual language instruction, and it allows for monitoring of progress and individualisation of language instruction. BeL (Borneo Learning) served as the reading platform, providing academic passages with cloud-based comprehension tasks and collaborative annotation tools for written corrective feedback. The platform's continuous assessment structure linking task completion to individual feedback cycles supported the iterative development of reading comprehension skills documented in the regression results. The reading domain's largest coefficient is consistent with findings from prior studies emphasizing the centrality of reading to language development and the effectiveness of personalized, feedback-rich digital environments. This finding adds weight to the studies of Holubnycha, et.al. (2019) which emphasized the need of mobile and cloud-based technologies in personalization, flexibility, and engaging language learning.

FlipGrid was used exclusively for speaking development, requiring students to submit weekly video responses, complete peer evaluations, and engage in asynchronous dialogue. The regression coefficient for Speaking ($B = 1.398$; standardized $\beta = .356$) was the smallest among the four domains, though it remained statistically significant ($p < .001$). This pattern is consistent with prior evidence that speaking is among the most cognitively demanding skills to develop within a short intervention period, particularly in an asynchronous, non-immersive format. The findings of this study are consistent with the findings of the study conducted by Chien, et.al (2020); Pirdayanti, et.al (2022), which explored the impact of video-based platform in the learning process of speaking and listening development.

Padlet was used for writing development through weekly argumentative paragraph and reflective essay tasks on a shared multimodal collaborative board, with structured peer and teacher feedback targeting coherence, vocabulary, and grammatical accuracy. The Writing coefficient ($B = 1.342$; $\beta = .367$) was the smallest unstandardized value but comparable in standardized form to

Listening (β = .371), indicating that Padlet-mediated written interaction supported skill development at a level approximately equal to the VoiceThread-mediated listening domain. This is consistent with evidence that collaborative digital writing platforms support a well-rounded approach to language development when structured around criterion-referenced feedback, Chamba & Chikusvura (2024) who has shown how using collaborative digital workspaces can have a positive impact on developing a well-rounded approach across all language domains.

VoiceThread is designed to enhance understanding and participation through voice, text and drawing features. Multimedia is versatile, can be manipulated, dynamic and interactive to satisfy different needs of learning and help master language; the comments are chronological. VoiceThread is multimodal, meaning that it can include all three modes of language (audio, visual, and text) to deepen students' understanding and retention of the language concepts presented, which then leads to an increase in language proficiency. This result is in line with Mayer (2024) Cognitive Theory of Multimedia Learning, which suggests that language concepts can be better comprehended and retained through using multiple sensory modalities.

The results of both the regression analysis and the platform-specific performance data give converging evidence of the model of the domain-platform alignment proposed in this study. The results of regression coefficients show that all these four domains of skills are significant and Reading has the highest unstandardized coefficient (B = 1.523; standardized β = .433) in predicting the Total Language Proficiency Score. The answers to the platform-level questions also indicate a development swing in each area that was grouped around the design of the platform. The results when combined reveal the benefits of using a multi-platform, skill differentiated approach to the implementation of one tool in the EFL higher education setting.

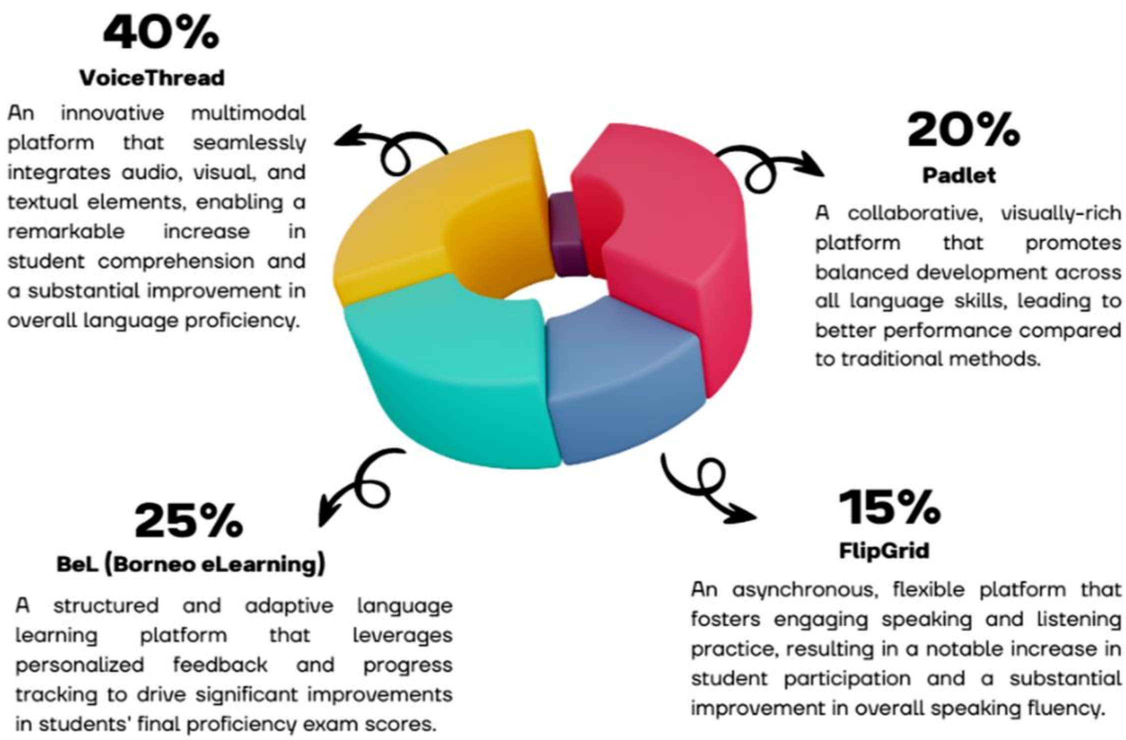

Figure 1. The Impact of Digital Learning Platforms

Figure 1 shows the domain specific regression coefficients developed through the analysis. Reading (BeL) has the highest unstandardized coefficient (B = 1.523) as it is the most influencing predictor in Total Language Proficiency Score. Overall, the standardized coefficient (β = .371 for the VoiceThread and β = .367 for the Padlet) associated with Listening (VoiceThread) and that associated with Writing (Padlet) did not differ significantly, suggesting that these two forms of representation (multimodal audio-visual and collaborative writing space) contributed at approximately the same level to the total level of proficiency. The smallest coefficient standardized was for Speaking (FlipGrid) (β = .356) and yet, this was statistically significant. All of the above indicate that differentiated instruction based on skill is needed and that more than one platform is required to serve the needs of the language domain, but also the needs of the platform.

4. DISCUSSION

Before discussing the substantive implications of the findings, an important interpretive boundary must be established. The regression coefficients reported in this study reflect the relative contribution of each skill domain score to the composite Total Language Proficiency Score within a sample exposed to domain-differentiated platform instruction. Because the intervention assigned each platform exclusively to one domain, the coefficient pattern is interpretively consistent with the hypothesis that each platform supported development in its target domain. However, the regression model does not include a control condition, does not directly compare pre-test to post-test change, and does not isolate the effect of each platform from other instructional variables. The current model primarily demonstrates domain-level proficiency patterns, not causally confirmed platform effects. Claims about which platform produced which gain are interpretively supported by the design but cannot be statistically confirmed from the present data alone. More direct evidence such as a randomized controlled design with platform-differentiated and platform-undifferentiated conditions would be required to make confident platform-specific causal claims.

With this interpretive boundary in mind, the ordering of regression coefficients (Reading > Listening ≈ Writing > Speaking) is consistent with sociocultural theory of language acquisition, which considers reading comprehension foundational to the development of other language skills (Baki, 2020; Nation, 2019). The near-equivalent Writing, Listening, and Speaking coefficients further suggest that the multi-platform, domain-differentiated design fostered comparatively balanced development across these three skills, in contrast to single-platform or single-skill studies where one domain typically dominates gains. These patterns provide preliminary support for the practical value of the domain-platform alignment model as a framework for making platform selection decisions in 3T-region EFL contexts (Alenezi, 2023; Marienko, et.al, 2020; Saykili, 2019).

The personalization and adaptive feedback features of BeL appear complementary to the regression model's emphasis on reading as the strongest domain predictor. In 3T-region contexts where authentic English input outside the classroom is limited, cloud-based reading tasks with written corrective feedback may provide a particularly valuable learning resource. The comparable regression weights for FlipGrid (Speaking) and VoiceThread (Listening) suggest that both video-based interaction and audio-visual annotation formats were associated with communicative skill development, though neither significantly outperformed the other in its contribution to overall proficiency. The collaborative structure of Padlet (Writing) and its cross-domain interaction between written expression and knowledge co-construction produced a domain contribution comparable to Listening, indicating that the platform's multimodal workspace supported meaningful written language development under structured conditions.

Speaking (FlipGrid) produced the smallest standardized coefficient ($\beta = .356$) among the four domains. This may reflect the inherently demanding nature of speaking skill development within a 12-week, asynchronous-only intervention, where production-oriented practice is constrained by the absence of real-time interactive feedback. It may also reflect a ceiling in the sensitivity of the speaking assessment instrument to gains within this population. Future studies should examine whether extending the intervention to 24 weeks or incorporating synchronous speaking tasks alongside asynchronous video interaction changes the relative contribution of the speaking domain.

This study has several limitations that constrain the interpretation and generalizability of findings. Most importantly, the use of domain subscores as predictors of a composite total score that is mathematically derived from the same subscores introduces a compositional constraint. From a mathematical standpoint, a high R^2 is expected in such a model and cannot be interpreted as evidence of predictive validity against an independent criterion. The regression coefficients are therefore best understood as estimates of the relative weight of each domain in the composite, under the specific conditions of this intervention, rather than as evidence of generalizable predictive relationships. Future studies should use an independently administered standardized test such as TOEFL or IELTS as the criterion variable to assess external validity. The single-institution design limits generalizability to other institutional contexts. The absence of a randomized control group means that observed patterns cannot be attributed causally to the platform-differentiated instruction without ruling out competing explanations. Future research should incorporate quasi-experimental designs with active control groups, longitudinal follow-up to assess retention of gains, and multi-site comparisons across 3T-region institutions in Indonesia (Guillén, et.al, 2021; Núñez, et.al, 2022).

5. CONCLUSION

This study examined the domain-specific alignment of mobile-cloud platforms to language skill domains and investigated whether domain-level contributions to total proficiency could be modelled differentially using multiple linear regression among Indonesian EFL university students in a 3T region. Four platforms were used BeL (reading), FlipGrid (speaking), Padlet (writing), and VoiceThread (listening) with each platform assigned exclusively to one skill domain across a 12-week intervention. A statistically significant regression model was obtained ($F=146.27$, $p < .001$; $R^2 = 0.921$), with all four domain predictors contributing significantly. Reading showed the largest domain contribution ($B= 1.523$; standardized $\beta = .433$), followed by listening ($\beta = .371$), writing ($\beta = .367$), and speaking ($\beta = .356$). The results are consistent with existing literature on the role of reading comprehension in integrated language development and support the hypothesis that domain-differentiated multi-platform instruction produces relatively balanced skill development across the remaining three domains.

Two points of contribution emerge from this study. First, the domain-platform alignment framework in which each platform is systematically assigned to one skill domain enables domain-level contributions to overall proficiency to be compared within a single study, providing a more structured basis for platform selection than general LMS effectiveness claims. Second, the finding that all four domain-platform pairings contributed significantly to total proficiency, with reading showing the strongest contribution and the other three domains contributing comparably, offers preliminary empirical support for balanced, domain-differentiated instruction in 3T-region EFL higher education. These contributions are preliminary and explicitly provisional: the regression model captures compositional domain relationships within the Total Score rather than independent predictive validity, and the absence of a randomized control group means that causal inferences about platform-specific gains cannot be drawn from the present design alone.

Directions for future research include longitudinal studies tracking proficiency development beyond 12 weeks; multi-site comparisons across 3T-region universities to assess the replicability and generalizability of the domain-coefficient pattern; quasi-experimental designs incorporating active control groups receiving undifferentiated platform instruction; and studies using an externally validated standardized test as the criterion variable to assess the external validity of the domain-platform alignment model. Research on teacher professional development for sustainable multi-platform implementation in low-resource settings is also warranted, as the practical value of the framework depends on educators having the capacity to deploy and manage multiple platforms coherently within a single course.

DECLARATION OF GENERATIVE AI

The authors further state that there was no involvement of generative artificial intelligence (AI) tools or AI-assisted technologies in writing, editing, analysis or preparation of this manuscript. The authors are the source of the material and guarantee the accuracy and integrity of the material in this study.

6. REFERENCES

Akram, H., Yingxiu, Y., Al-Adwan, A. S., & Alkhalifah, A. (2021). Technology Integration in Higher Education During COVID-19: An Assessment of Online Teaching Competencies Through Technological Pedagogical Content Knowledge Model. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.736522

Alenezi, M. (2023). Digital Learning and Digital Institution in Higher Education. *Education Sciences*, *13*(1), 88. https://doi.org/10.3390/educsci13010088

Amhag, L., Hellström, L., & Stigmar, M. (2019). Teacher Educators' Use of Digital Tools and Needs for Digital Competence in Higher Education. *Journal of Digital Learning in Teacher Education*, *35*(4), 203–220. https://doi.org/10.1080/21532974.2019.1646169

Andringa, S., & Godfroid, A. (2020). Sampling Bias and the Problem of Generalizability in Applied Linguistics. *Annual Review of Applied Linguistics*, *40*, 134–142. https://doi.org/10.1017/S0267190520000033

Asratie, M. G., Wale, B. D., & Aylet, Y. T. (2023). Effects of using educational technology tools to enhance EFL students' speaking performance. *Education and Information Technologies*, *28*(8), 10031–10051. https://doi.org/10.1007/s10639-022-11562-y

Baki, Y. (2020). The Effect of Critical Reading Skills on the Evaluation Skills of the Creative Reading Process. *Eurasian Journal of Educational Research*, *20*(88), 1–26. https://doi.org/10.14689/ejer.2020.88.9

Chamba, L. T., & Chikusvura, N. (2024). Future-proofing quality education using integrated assessment systems. *Quality Education for All*, *1*(1), 240–255. https://doi.org/10.1108/QEA-11-2023-0014

Chen, C.-M., Li, M.-C., & Lin, M.-F. (2022). The effects of video-annotated learning and reviewing system with vocabulary learning mechanism on English listening comprehension and technology acceptance. *Computer Assisted Language Learning*, *35*(7), 1557–1593. https://doi.org/10.1080/09588221.2020.1825093

Chien, S.-Y., Hwang, G.-J., & Jong, M. S.-Y. (2020). Effects of peer assessment within the context of spherical video-based virtual reality on EFL students' English-Speaking performance and learning perceptions. *Computers & Education*, *146*, 103751. https://doi.org/10.1016/j.compedu.2019.103751

Dunn, T. J., & Kennedy, M. (2019). Technology Enhanced Learning in higher education; motivations, engagement and academic achievement. *Computers & Education*, *137*, 104–113. https://doi.org/10.1016/j.compedu.2019.04.004

Dzvinchuk, D., Radchenko, O., Kachmar, O., Myskiv, I., & Dolinska, N. (2020). Analysis of

Platforms and Tools of Open Study in the Conditions of Postmodern Education. *Revista Romaneasca Pentru Educatie Multidimensionala*, *12*(3), 125–143. https://doi.org/10.18662/rrem/12.3/313

Estrada, P., Wang, H., & Farkas, T. (2020). Elementary English Learner Classroom Composition and Academic Achievement: The Role of Classroom-Level Segregation, Number of English Proficiency Levels, and Opportunity to Learn. *American Educational Research Journal*, *57*(4), 1791–1836. https://doi.org/10.3102/0002831219887137

Groening, C., & Binnewies, C. (2019). "Achievement unlocked!" - The impact of digital achievements as a gamification element on motivation and performance. *Computers in Human Behavior*, *97*, 151–166. https://doi.org/10.1016/j.chb.2019.02.026

Guillén-Gámez, F. D., Mayorga-Fernández, M. J., Bravo-Agapito, J., & Escribano-Ortiz, D. (2021). Analysis of Teachers' Pedagogical Digital Competence: Identification of Factors Predicting Their Acquisition. *Technology, Knowledge and Learning*, *26*(3), 481–498. https://doi.org/10.1007/s10758-019-09432-7

Gullifer, J. W., Kousaie, S., Gilbert, A. C., Grant, A., Giroud, N., Coulter, K., … Titone, D. (2021). Bilingual language experience as a multidimensional spectrum: Associations with objective and subjective language proficiency. *Applied Psycholinguistics*, *42*(2), 245–278. https://doi.org/10.1017/S0142716420000521

Holubnycha, L., Kostikova, I., Kravchenko, H., Simonok, V., & Serheieva, H. (2019). Cloud Computing for University Students' Language Learning. *Revista Romaneasca Pentru Educatie Multidimensionala*, 55–69. https://doi.org/10.18662/rrem/157

Huang, A. Y. Q., Lu, O. H. T., & Yang, S. J. H. (2023). Effects of artificial Intelligence–Enabled personalized recommendations on learners' learning engagement, motivation, and outcomes in a flipped classroom. *Computers & Education*, *194*, 104684. https://doi.org/10.1016/j.compedu.2022.104684

Kilinç, E., & Tarman, B. (2022). Citizenship types, social media use and speaking a foreign language as predictors of global competence. *Citizenship Teaching & Learning*, *17*(1), 49–62. https://doi.org/10.1386/ctl_00081_1

Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Liu, S., & Da, J. (2021). Technology in Chinese Language Teaching. In *The Palgrave*

Handbook of Chinese Language Studies (pp. 1–41). Singapore: Springer Singapore. https://doi.org/10.1007/978-981-13-6844-8_3-1

Marienko, M., Nosenko, Y., Sukhikh, A., Tataurov, V., & Shyshkina, M. (2020). Personalization of learning through adaptive technologies in the context of sustainable development of teachers' education. *E3S Web of Conferences*, *166*, 10015. https://doi.org/10.1051/e3sconf/202016610015

Mayer, R. E. (2024). The Past, Present, and Future of the Cognitive Theory of Multimedia Learning. *Educational Psychology Review*, *36*(1), 8. https://doi.org/10.1007/s10648-023-09842-1

Metsäpelto, R.-L., Poikkeus, A.-M., Heikkilä, M., Husu, J., Laine, A., Lappalainen, K., … Suvilehto, P. (2022). A multidimensional adapted process model of teaching. *Educational Assessment, Evaluation and Accountability*, *34*(2), 143–172. https://doi.org/10.1007/s11092-021-09373-9

Mohamed, S., & Adnan, F. (2020). Feedback in Computer-Assisted Language Learning: A Meta-Analysis. *Tesl-Ej*, *24*(2), n2.

Nation, K. (2019). Children's reading difficulties, language, and reflections on the simple view of reading. *Australian Journal of Learning Difficulties*, *24*(1), 47–73. https://doi.org/10.1080/19404158.2019.1609272

Núñez-Canal, M., de Obesso, M. de las M., & Pérez-Rivero, C. A. (2022). New challenges in higher education: A study of the digital competence of educators in Covid times. *Technological Forecasting and Social Change*, *174*, 121270. https://doi.org/10.1016/j.techfore.2021.121270

Pirdayanti, N. P. A., Nitiasih, P. K., Budiarta, L. G. R., & Adnyayanti, N. L. P. E. (2022). The Impact of Videos Based Discovery Learning Towards Young Learners' Speaking Skill During Pandemic Covid-19. *Jurnal Pedagogi Dan Pembelajaran*, *5*(2), 285–292. https://doi.org/10.23887/jp2.v5i2.45556

Rahayu, Weda, S., Muliati, & De Vega, N. (2024). Artificial Intelligence in writing instruction: A self-determination theory perspective. *XLinguae*, *17*(1), 234–244. https://doi.org/10.18355/XL.2024.17.01.16

Rahimi, R. A., & Oh, G. S. (2024). Rethinking the role of educators in the 21st century: navigating globalization, technology, and pandemics. *Journal of Marketing Analytics*, *12*(2), 182–197. https://doi.org/10.1057/s41270-024-00303-4

Saubern, R., Urbach, D., Koehler, M., & Phillips, M. (2020). Describing increasing proficiency in teachers' knowledge of the effective use of digital technology. *Computers & Education*, *147*, 103784. https://doi.org/10.1016/j.compedu.2019.103784

Saykili, A. (2019). Higher Education in The Digital Age: The Impact of Digital Connective Technologies. *Journal of Educational Technology and Online Learning*, *2*(1), 1–15. https://doi.org/10.31681/jetol.516971

Selfa-Sastre, M., Pifarré, M., Cujba, A., Cutillas, L., & Falguera, E. (2022). The Role of Digital Technologies to Promote Collaborative Creativity in Language Education. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.828981

Septiyana, L., Safitri, A., & Aminatun, D. (2021). The Correlation Between EFL Learners Cohesion and Their Reading Comprehension. *Journal of Research on Language Education*, *2*(2), 68. https://doi.org/10.33365/jorle.v2i2.1154

Su, F., & Zou, D. (2022). Technology-enhanced collaborative language learning: theoretical foundations, technologies, and implications. *Computer Assisted Language Learning*, *35*(8), 1754–1788. https://doi.org/10.1080/09588221.2020.1831545

Verma, J. P., & G. Abdel-Salam, A. (2019). *Testing Statistical Assumptions in Research*. Wiley. https://doi.org/10.1002/9781119528388

Walker, R. (2020). Facilitating Active Learning Opportunities for Students Through the Use of Technology-Enhanced Learning Tools: The Case for Pedagogic Innovation and Change. In *Teaching Learning and New Technologies in Higher Education* (pp. 117–133). Singapore: Springer Singapore. https://doi.org/10.1007/978-981-15-4847-5_9