**International Journal of Educational Qualitative Quantitative Research (IJE-QQR)**
ISSN: 2962-9918 (Online)
Vol. 5, No. 1, pp. 81-91



# Enhancing IELTS Speaking Preparation: The Development of a Self-Paced Learning Model Integrating Edpuzzle and Authentic Materials

**Kurnihayati Kurnihayati[1], Endang Siti Nurkholidah[2], Affan Phatthanawiwat[3]**
[1,2]English Langauge and Education Study Program, Faculty of Teacher Training and Education, Universitas Sindang Kasih Majalengka, Indonesia
[3]Thammasat University, Bangkok, Thailand

**Article Info**

**Article history:**
Received: Feb 15, 2026
Revised: June 11, 2026
Accepted: June 18, 2026

DOI: 10.58418/ijeqqr.v5i1.200

**How to cite this article:**
Kurnihayati, K., Nurkholidah, E. S., & Phatthanawiwat, A. (2026). Enhancing IELTS Speaking Preparation: The Development of a Self-Paced Learning Model Integrating Edpuzzle and Authentic Materials. *International Journal of Educational Qualitative Quantitative Research*, *5*(1), 81–91. https://doi.org/10.58418/ijeqqr.v5i1.200

**Read online:**

*Scan this QR code with your smart phone or mobile device to read online.*

**ABSTRACT**

High-stakes oral language assessments often present significant challenges for English as a Foreign Language (EFL) learners due to a lack of autonomous, context-rich speaking environments that support spontaneous lexical and grammatical development. To address this limitation, the present study developed and evaluated a self-paced learning model that integrates the Edpuzzle platform with authentic instructional materials specifically designed for International English Language Testing System (IELTS) Speaking test preparation. Operating within a Research and Development (R&D) design, this project applied the 4D framework across its four successive phases of define, design, develop, and disseminate to pilot the model with ten English Education majors from three Indonesian universities. Empirical data were collected and triangulated using evaluation questionnaires, semi-structured interviews, and asynchronous user engagement logs extracted directly from the platform. The findings demonstrated noticeable performance enhancements across key evaluation criteria. Quantitative assessments indicated highly favorable participant reception regarding pronunciation tasks (M = 4.8) and contextual real-world relevance (M = 4.7), whereas real-time grammatical accuracy presented a moderate ongoing challenge for learners (M = 3.0). Qualitative responses corroborated these trends, revealing that the interactive multimedia architecture reinforced learner autonomy and reflective language use, which mitigated minor technical glitches and dialectal challenges. Conceptually, this research bridges a critical gap by shifting technology-mediated authentic learning out of traditional teacher-centered classrooms into a fully self-directed environment. Practically, it offers an accessible, scalable pedagogical framework for autonomous test preparation, thereby advancing multimedia instructional design within global EFL frameworks.

**Keywords:** Authentic Materials, Computer-Assisted Language Learning, Edpuzzle, IELTS speaking, Self-Paced Learning Model



***Corresponding Author:***
Kurnihayati Kurnihayati
English Langauge and Education Study Program, Faculty of Teacher Training and Education, Universitas Sindang Kasih Majalengka, Indonesia
Email: kurnihayati@gmail.com

## 1. INTRODUCTION

In the contemporary globalized landscape, professional mobility and academic advancement heavily rely on standardized English language certifications, with the International English Language Testing System (IELTS) serving as a primary benchmark (Coney & Isbell, 2026; Maryam et al., 2025; Read, 2022). Consequently, achieving oral proficiency for high-stakes tests has transitioned from an optional skill to an

institutional imperative across corporate and academic sectors worldwide. Within expanding economies like Indonesia, this demand is mirrored by state-owned enterprises and higher education institutions implementing strict communicative competence thresholds for employment and graduation. Despite its critical importance, mastering the IELTS Speaking component remains an elusive milestone for many English as a Foreign Language (EFL) learners. Data from previous studies indicate that students continuously encounter deep-seated obstacles, ranging from lexical deficiencies and grammatical inaccuracies (Octaberlina et al., 2022; Taye & Mengesha, 2024) to severe affective barriers such as test anxiety (Akram & Abdelrady, 2023) and communicative apprehension (Kuluşaklı & Genç, 2024).

The pedagogical urgency to solve these oral deficiencies is further compounded by socioeconomic stratification and classroom limitations. In non-metropolitan regions, formal IELTS training centers impose prohibitive financial constraints that make structured preparation inaccessible for the general student population. Furthermore, conventional classroom settings typically fail to bridge this gap due to rigid instructional durations, minimal exposure to native speech inputs, and less interactive teaching methods that stifle spontaneous communication (Darmawansah et al., 2025; Mahdi, 2022; Ostovar-Namaghi et al., 2024). To mitigate these logistical and financial limitations, shifting toward an autonomous, technology-mediated framework through a self-paced learning model has emerged as a critical necessity (Mekheimer, 2025; Nikolopoulou, 2023). This independent model allows learners to regulate their own time, pace, and strategies, which directly addresses the lack of optimal practice opportunities found in traditional educational structures.

To ensure the efficacy of independent oral preparation, the integration of authentic learning materials is highly urgent. Materials such as global news broadcasts, contemporary articles, digital advertisements, and unscripted interviews expose learners to genuine language ecosystems outside the artificial constraints of traditional textbooks, thereby enhancing lexical resource and pronunciation. However, recent empirical evidence warns that raw authentic data can induce severe cognitive overload if deployed without structured pedagogical scaffolding (Alamri, 2025; Gkintoni et al., 2025). Therefore, a critical urgency exists to pair authentic inputs with an interactive platform like Edpuzzle. Edpuzzle serves as a vital pedagogical scaffolding tool that transforms passive multimedia consumption into an active learning experience through embedded formative checks, asynchronous feedback loops, and structured speaking stages (Dmitrenko & Panchenko, 2025; Setiawati et al., 2025). Synthesizing authentic digital inputs within Edpuzzle allows learners to navigate progressive speaking tasks, from imitative to extensive production, while effectively managing their cognitive load.

Rooted in the Cognitive Theory of Multimedia Learning (Mayer, 2024; Twabu, 2025), the pedagogical synthesis of interactive video platforms and authentic inputs relies on the balanced activation of dual-coding processing channels. When EFL learners engage with authentic video materials, they process simultaneous visual and auditory streams, which can easily trigger extraneous cognitive load if left unmanaged. According to Zahara (2025), Edpuzzle mitigates this cognitive risk by allowing instructional designers to segment complex native speech into digestible chunks, embed interactive prompts, and provide immediate reflective pauses. This platform-mediated scaffolding transforms passive listening into an active cognitive reconstruction process, which is pivotal for transferring input comprehension into oral output production. By utilizing structured digital prompts to segment authentic language, the instructional design successfully balances working memory constraints, thereby optimizing the cognitive pathways required for spontaneous fluency and phonetic accuracy during high-stakes test preparation.

Recent literature has extensively explored the separate parameters of speaking instruction, platform usage, and learner autonomy. For instance, Nurdiana et al. (2025) investigated self-directed learning strategies in oral communication but stopped short of proposing structured technological models designed for standardized test parameters. Concurrently, Lestari et al. (2023) developed Edpuzzle-based media for listening and speaking classes, yet their framework remained localized within teacher-controlled, synchronous classroom environments. While Thai & Nguyen (2022) highlighted the communicative benefits of authentic digital texts, their implementation was confined to conventional secondary classrooms without exploiting the self-regulated affordances of autonomous learning platforms. More recently, research has underscored how authentic oral production within digital ecosystems enhances fluency and mitigates technological disparity (Hu et al., 2025; Nieminen et al., 2023; Oshimeje & Flores Barahona, 2025). However, a significant empirical gap persists regarding the systematic convergence of these variables into a unified instructional design.

The novelty of this study lies in the synergistic integration of Edpuzzle architecture and authentic materials into a fully autonomous, self-paced instructional model specifically validated for high-stakes IELTS Speaking preparation. Unlike previous works that treat technology-mediated authentic learning as a teacher-led classroom supplement, this research introduces a standalone, validated framework that shifts the locus of control entirely to the independent learner. By framing this intervention within a representative regional Indonesian higher education context through a Research and Development (R&D) design, this study utilizes the 4D model encompassing the sequential phases of define, design, develop, and disseminate. Specifically, this paper addresses three interconnected research questions regarding how the self-paced learning model integrating Edpuzzle and authentic materials is systematically developed, what actual

learning experiences students encounter within this self-paced digital environment, and what logistical or cognitive challenges they face during autonomous implementation.

## 2. METHOD

### 2.1. Research Design

This study adopted a Research and Development (R&D) design based on the 4D development model, which encompasses the four sequential phases of define, design, develop, and disseminate. This systematic framework was selected to ensure a disciplined, iterative process for conceptualizing, validating, and evaluating the self-paced learning model. The operationalization of the 4D framework progressed through distinct stages starting from a front-end analysis to examine learning difficulties and technological readiness in International English Language Testing System (IELTS) Speaking preparation. The design and develop phases involved blueprinting the instructional materials, generating the prototype on the Edpuzzle platform, and conducting expert validation, while the disseminate phase was executed via a restricted cross-institutional dissemination to evaluate practical adaptability across multiple higher education settings.

### 2.2. Context and Participants

The field trial was executed over a one-month period within the English Education departments of three higher education institutions in Indonesia: Universitas Islam Negeri Siber Syekh Nurjati Cirebon, Universitas Swadaya Gunung Jati, and Universitas Muhammadiyah Cirebon. A cohort of ten undergraduate students was chosen through a purposive sampling technique to serve as the primary user group for this operational testing. The inclusion criteria dictated that participants must be actively preparing for high-stakes standardized English assessments, have regular access to digital devices, and exhibit high engagement in self-directed learning frameworks. Employing a small yet multi-institutional cohort allowed the study to observe the ecological validity and cross-contextual adaptability of the autonomous platform during the restricted dissemination phase.

### 2.3. Research Instruments and Validation Framework

Data collection relied on a mixed-method triangulation design to capture both quantitative perceptions and qualitative behavioral insights. Three main instruments were utilized to gather comprehensive data, namely a perception questionnaire, a semi-structured interview protocol, and an asynchronous platform observation log. To ensure measurement rigor, the questionnaire items were adapted from established computer-assisted language learning frameworks and modified to align with official IELTS Speaking band descriptors. The structural blueprint of this quantitative instrument is explicitly outlined in Table 1 to show the alignment between survey items and evaluated skills. Lexical resource adapted from Asratie et al. (2023); Grammatical range and accuracy adapted from Liu et al. (2026); Material authenticity adapted from Alamri (2025); Pronunciation adapted from Bogach et al. (2021); Fluency and coherence adapted from Wilang et al. (2026).

**Table 1.** Structural Blueprint of the Perception Questionnaire

| Evaluation Indicator | Survey Item Numbers | Focus of Assessment |
|---|---|---|
| Lexical Resource | 1, 2 | Range, variety, and contextual appropriateness of vocabulary |
| Grammatical Range and Accuracy | 3, 4 | Morphosyntactic complexity and systematic error control |
| Material Authenticity | 5, 6 | Real-world linguistic relevance and cultural exposure |
| Pronunciation | 7, 8 | Phonetic clarity, intonation patterns, and speech rhythm |
| Fluency and Coherence | 9, 10 | Spontaneous speech rate, continuity, and logical linking |

The structural composition presented in Table 1 functions as the baseline metric to quantify user satisfaction regarding individual linguistic components. By renumbering and compartmentalizing the survey items into these ten focused indicators, the instrument allows the researcher to isolate exactly which speaking competencies receive the highest or lowest support from the multimedia architecture. This structural layout ensures that the quantitative data directly correspond to the official evaluative parameters of high-stakes oral testing. To complement these psychometric dimensions, qualitative tools were implemented to capture the descriptive nuances of the user experience. The conceptual trajectory of the qualitative investigation is organized within a separate, structured protocol as detailed in Table 2. Operational and technical challenges adapted from Kolapo et al. (2025); Autonomous learning experiences adapted from Hussain et al. (2025); Perceived pedagogical impact adapted from Ferrari (2023).

**Table 2.** Blueprint of the Semi-Structured Interview Protocol

| Investigation Domain | Interview Items | Operational Focus |
|---|---|---|
| Operational and Technical Challenges | 1, 2 | Interface friction, connectivity issues, and device bugs |
| Autonomous Learning Experiences | 3 | Meta-cognitive monitoring and self-pacing habits |
| Perceived Pedagogical Impact | 4, 5 | Confidence growth and skill transfer to actual tasks |

The operational focus established in Table 2 serves as a guide to explore the internal cognitive and external logistical realities encountered by the student cohort. By targeting specific domains using a dedicated five-item qualitative protocol, the interview questions protect the study from anecdotal bias. This alignment ensures the qualitative narratives directly corroborate or explain the trends discovered in the statistical evaluations. To interpret the quantitative scores retrieved from the field trials, a formal mathematical conversion framework was established. The categorical benchmarks for interpreting the mean scores of both product validation and user practicality are systematically presented in Table 3.

**Table 3.** Quantitative Interpretation and Decision Criteria

| Mean Score Interval | Validity Level | Practicality Status | Operational Meaning |
|---|---|---|---|
| 4.20 to 5.00 | Highly Valid | Highly Practical | Implementable without any structural revisions |
| 3.40 to 4.19 | Valid | Practical | Implementable with minor multimedia enhancements |
| 2.60 to 3.39 | Moderately Valid | Moderately Practical | Requires selective structural and content changes |
| 1.00 to 2.59 | Invalid | Impractical | Not implementable; requires complete redesign |

The numerical boundaries specified in Table 3 provide the analytical discipline required to evaluate the product objectively. By pinning the calculated mean scores to predefined validity levels and operational meanings, the framework removes subjectivity when translating student feedback into design modifications. This ensures that any conclusions drawn regarding the efficacy of the Edpuzzle module are mathematically grounded and methodologically sound.

### 2.4. Data Collection Procedures

The data collection procedures followed a systematic timeline to ensure high administrative control. After completing the module on the Edpuzzle platform, participants received the online questionnaire via Google Forms to record their psychometric evaluations. Following this, semi-structured online interviews were conducted with selected participants to capture deep-seated cognitive reflections. Concurrently, behavioral data were extracted from the backend analytics of the Edpuzzle platform to observe real-time interactions asynchronously. This asynchronous observation allowed the researcher to track non-intrusive operational metrics, such as individual video completion rates, exact viewing durations, and the frequency of replaying video segments. Ethical compliance was strictly maintained by securing signed informed consent from all participants and ensuring complete data anonymization.

### 2.5. Data Analysis and Rigor

Quantitative data derived from the survey were analyzed using descriptive statistics to calculate the Mean score for each target indicator, establishing the baseline consensus of user reception. Qualitative insights from the student interviews and platform analytics were transcribed, managed, and analyzed using a rigorous thematic analysis approach. This involved a multi-stage process of data familiarization, initial code generation, and inductive theme clustering to isolate core themes related to fluency, authenticity, and technical barriers. To ensure qualitative trustworthiness, investigator triangulation was used by cross-referencing interview transcripts with actual platform logs. Additionally, member-checking was conducted by returning the summarized transcript findings to the interviewees to verify interpretive accuracy and eliminate subjective researcher bias.

## 3. RESULTS

### 3.1. Development and Validation of the Self-Paced Learning Model

The development of the self-paced learning model integrating the Edpuzzle platform and authentic instructional materials progressed systematically through the four stages of the 4D framework. During the initial define phase, a front-end diagnostic analysis was carried out across the participating higher education institutions to map out the foundational difficulties encountered by undergraduate EFL learners. The

diagnostic findings revealed that the primary instructional bottlenecks centered on constrained lexical resources, morphosyntactic inaccuracies, low speaking confidence, and poor spontaneous idea generation. Pronunciation difficulties were also noted by the learners but were evaluated as significantly less critical than the structural and affective barriers. These collective field insights established a pressing pedagogical urgency to build a flexible, technology-mediated learning model that simultaneously targets linguistic mechanics and psychological readiness for the high-stakes IELTS Speaking assessment.

Moving into the design phase, the curriculum architecture was established by curating a diverse repository of real-world materials including global news broadcasts, unscripted native conversations, and context-rich speaking tasks. This curated digital content was structured into a progressive pedagogical blueprint featuring four distinct learning tiers consisting of imitative, intensive, responsive, and extensive language production tasks. This tiered arrangement was designed to provide scaffolded practice that prevents immediate cognitive overload while systematically building oral proficiency. The resulting blueprint served as the initial media prototype before the project advanced into formal expert assessment and technical refinement.

The develop phase focused extensively on validating the structural soundness, curricular alignment, and technical functionality of the instructional prototype. The review panel was comprised of three external experts, including two senior professors in English applied linguistics who examined content validity, alongside one specialist in educational technology who evaluated the multimedia architecture. Each evaluator evaluated the interactive platform independently using a standardized instrument structured on a five-point Likert scale. To establish complete empirical transparency regarding the structural integrity of the R&D product, the quantitative marks provided by the validation panel are explicitly compiled in Table 4.

**Table 4.** Quantitative Results of Expert Validation Trials

| Validator Profile | Evaluation Focus Domain | Assigned Mean Score | Categorical Validity Level |
|---|---|---|---|
| External Expert 1 (Senior Professor) | Content Accuracy and IELTS Rubric Alignment | 4.60 | Highly Valid |
| External Expert 2 (Senior Professor) | Curricular Flow and Linguistic Appropriateness | 4.40 | Highly Valid |
| External Expert 3 (Technology Specialist) | Multimedia Architecture and Usability Design | 4.50 | Highly Valid |

The quantitative metrics displayed in Table 4 demonstrate that the developed prototype achieved a highly valid status across all evaluation channels, directly satisfying the baseline criteria for classroom implementation. In addition to assigning positive numerical scores, the validation panel provided critical qualitative recommendations aimed at balancing working memory demands in line with multimedia learning principles. The language experts recommended integrating textual summaries for rapid native speech segments, while the instructional technologist suggested adjusting the background audio tracks to enhance phonetic clarity. The precise formative modifications executed during this iterative development cycle are chronologically organized and displayed in Table 5.

**Table 5.** Formative Instructional Revisions Based on Qualitative Expert Feedback

| Evaluation Source | Identified Structural Critique | Executed Formative Revision |
|---|---|---|
| Linguistics Panel | Rapid native speech rates induce cognitive overload | Integrated pop-up text summaries and strategic pauses |
| Multimedia Specialist | Background audio tracks overlap with target pronunciation | Calibrated audio levels and eliminated ambient noises |

The operational execution of the structural adjustments outlined in Table 5 directly optimized the instructional flow and linguistic appropriateness of the digital module prior to field testing. The final disseminate phase was subsequently executed through a restricted cross-institutional dissemination involving ten English Education majors across three distinct universities in Cirebon, Indonesia. Post-implementation surveys indicated that 100% of the student cohort agreed that the self-paced Edpuzzle model successfully facilitated their autonomous preparation for the IELTS Speaking test. Qualitative interviews corroborated these quantitative trends, with participants highlighting that the capacity to replay video clips at their own convenience fostered deep self-monitoring habits. This high level of user autonomy and platform flexibility is clearly demonstrated in the reflective statements volunteered by the student cohort.

*"Everything is complete; from vocab, grammar, pronunciation, to speaking. I can learn and practice by myself without needing a classroom."* (Participant 1)

*"I can use this learning media for self-study because the videos are interactive and easy to understand. The materials also help improve my speaking skills without having to constantly depend on the teacher."* (Participant 2)

The remaining members of the student cohort expressed similar praise for the interactive multimedia format, noting that the combination of authentic native models and embedded questions helped lower test-related anxiety. These convergent findings confirm that the final R&D product successfully transitions high-stakes test preparation into an accessible, flexible, and learner-centered digital space. Consequently, the model establishes a robust framework for self-regulated language acquisition without requiring constant, face-to-face academic supervision.

### 3.2. Evaluation of Students' Learning Experiences

The quantitative evaluation regarding students' learning experiences yielded highly favorable scores across the majority of the evaluated speaking sub-skills. The psychometric datasets captured from the five-point Likert scale survey reflected a strong student consensus regarding the practical utility of the platform features. To ensure comprehensive reporting and full empirical transparency, the detailed statistical distribution including the computed Mean and Standard Deviation scores for each evaluation descriptor is systematically compiled in Table 6.

**Table 6**. Descriptive Statistical Distribution of Students' Learning Experiences

| Evaluation Indicator | Targeted Pedagogical Sub-Aspect | Mean (M) | Standard Deviation (SD) |
|---|---|---|---|
| Pronunciation | Word repetition and phonetic clarity | 4.80 | 0.42 |
| Material Authenticity | Real-life communicative relevance | 4.70 | 0.48 |
| Lexical Resource | Vocabulary sufficiency within tasks | 4.50 | 0.53 |
| Lexical Resource | Vocabulary confidence and usage | 4.40 | 0.61 |
| Fluency and Coherence | Structured task sequencing | 4.40 | 0.52 |
| Fluency and Coherence | Task-driven fluency facilitation | 4.40 | 0.52 |
| Grammatical Range | Conceptual grammar understanding | 4.10 | 0.67 |
| Grammatical Accuracy | Spontaneous grammar application | 3.00 | 1.05 |

The descriptive data presented in Table 6 indicate that pronunciation tasks achieved the highest statistical rating (M = 4.80, SD = 0.42), which validates the effectiveness of word-repetition exercises driven by authentic native models. Strong positive evaluations were also recorded for real-world material relevance (M = 4.70, SD = 0.48) and vocabulary confidence (M = 4.40, SD = 0.61), showcasing a substantial enrichment within the learners' lexical resource domain. In terms of speech continuity, the structured sequencing of tasks (M = 4.40, SD = 0.52) successfully guided students to organize oral arguments naturally and maintain smoother speech over time. Conversely, a prominent developmental challenge was identified in spontaneous grammatical accuracy, which received the lowest evaluation score (M = 3.00, SD = 1.05). This specific statistical trend reveals that while the participants could conceptually comprehend complex morphosyntactic structures (M = 4.10, SD = 0.67), executing those rules accurately during real-time speech production remains a major cognitive barrier that requires additional scaffolded practice.

### 3.3. Structural Analysis of Logistical and Cognitive Challenges

Thematic analysis of the qualitative data streams isolated several operational and cognitive friction points encountered by the student cohort during their self-directed study blocks. These experiential challenges were systematically categorized into technical barriers, cognitive-linguistic difficulties, and future user-experience optimization suggestions. The exact typology, description, and recurrence frequencies of these identified user challenges are organized and presented in Table 7.

**Table 7.** Typology and Frequency of Encountered Learning Challenges

| Challenge Dimension | Core Identified Issues within Platform | Frequency Count (N = 10) |
|---|---|---|
| Technical Barriers | Unstable internet connectivity and streaming buffering | 8 |
| Technical Barriers | Operating system and iOS device browser incompatibility | 5 |
| Technical Barriers | Audio recording errors and background microphone noise | 4 |
| Cognitive-Linguistic | Rapid native speech rates and unfamiliar British dialects | 6 |
| Cognitive-Linguistic | Rapid instruction recall and performance anxiety | 3 |
| System Optimization | Lack of interactive subtitles and cue card guides | 7 |

The empirical parameters detailed in Table 7 identify unstable internet connectivity as the most pervasive technical obstacle, affecting eight out of the ten participants due to local bandwidth drops that disrupted video streaming. Device operating system compatibility also induced noticeable system friction, particularly for users accessing the module via iOS devices where the native audio controls required multiple browser refreshes to playback correctly. This specific technical glitch occasionally interrupted student concentration and delayed task completion, as explicitly reported by one of the affected learners during the post-trial evaluations.

*"When I opened it on iOS, the audio did not appear, so I had to refresh the page several times. It worked much better when I eventually switched to an Android device."* (Participant 3)

Linguistic challenges centered primarily on the rapid speech rates and dialectal variations embedded within the authentic materials, with six participants reporting initial difficulties interpreting unfamiliar British accents. Additionally, three students experienced mild performance anxiety when prompted to record their voice responses, which temporarily disrupted speech continuity and logical linking. To optimize future platform design, seven participants requested the addition of permanent subtitles for fast-paced native dialogues alongside shorter, segmented video modules. This optimization trajectory was summarized effectively by a student who emphasized the value of adding localized scaffolding items to the multimedia interface.

*"The video module is already good and easy to follow, but adding synchronized subtitles and structural examples for the speaking cue cards would make it even more useful for independent study."* (Participant 4)

Despite encountering these technical and linguistic friction points, the student cohort successfully managed their individual study goals by replaying complex sections or switching to alternative hardware. The overall field results indicate that the self-paced multimedia model remained highly functional, serving as a practical tool for independent IELTS Speaking preparation. Therefore, these operational outcomes confirm that the platform is robust enough to sustain independent language acquisition despite localized infrastructural limitations.

## 4. DISCUSSION

### 4.1. Pedagogical Validity and Multimedia Integration in Model Development

The systematic execution of the 4D framework provided a disciplined pathway for developing an instructional model that is both pedagogically valid and contextually grounded. By anchoring the baseline development in a front-end diagnostic analysis during the define phase, the resulting prototype directly targeted the documented vulnerabilities of undergraduate EFL learners, such as limited vocabulary range and low speaking confidence. This targeted engineering confirms the assertions of Blhaj et al. (2025) and Csapó & Molnár (2019), who argued that development models must prioritize localized diagnostic data to ensure the ecological validity of digital learning products. Furthermore, transferring the instructional control to an asynchronous space via Edpuzzle allowed learners to independently regulate their learning pace, select strategic review points, and manage their individual timelines. This operational environment successfully satisfies the classic self-directed learning tenets established by Ahmad (2024), Zhang & Dong (2024), and Kharroubi & ElMediouni (2024), which emphasize that adult learners thrive when granted meta-cognitive control over their instructional trajectories.

The inclusion of authentic video texts structured into a progressive hierarchy from imitative to extensive production stages offered an explicit scaffolding system. This structured sequencing enabled students to manage their cognitive capacity while transitioning from controlled phrase repetition to complex independent speaking tasks. Expert feedback during the development phase highlighted the necessity of aligning the platform interface with Cognitive Theory of Multimedia Learning (Marougkas et al., 2023; Mayer, 2024; Twabu, 2025) to prevent split-attention effects. Integrating interactive textual summaries and removing overlapping background audio elements successfully balanced the visual and auditory processing channels of the participants. Consequently, this structural alignment minimized extraneous cognitive load, thereby creating a highly supportive multimedia environment where learners could systematically internalize authentic oral communication patterns.

### 4.2. Deconstructing the Asymmetry in Students' Oral Learning Experiences

The statistical results extracted from the field trials exposed a compelling developmental asymmetry between the participants' phonetic performance and their spontaneous grammatical application. The high evaluation score recorded for pronunciation mastery ($M = 4.80$, $SD = 0.42$) demonstrates that constant, self-paced exposure to native acoustic models within authentic video clips effectively accelerates phonetic calibration. This rapid advancement aligns with the oral automaticity principles postulated by Ochoa (2022) and Kakitani (2025), which indicate that non-intrusive task repetition within a digitized space lowers the affective barriers to correct phonetic production. These dynamics are further substantiated by Mahdi (2022) and Alam (2025) who documented that integrating authentic communicative multimedia interactions

significantly optimizes the articulatory accuracy and oral confidence of EFL students preparing for high-stakes standardized assessments. Genuine linguistic exposure helps students mimic natural speech rhythms more effectively than traditional, teacher-led drills.

In sharp contrast to this phonetic growth, the prominent discrepancy between conceptual grammar understanding (M = 4.10, SD = 0.67) and spontaneous grammatical accuracy (M = 3.00, SD = 1.05) represents a critical pedagogical friction point. This clear developmental gap reveals that possessing a passive, receptive mastery of morphosyntactic rules does not automatically translate into real-time productive competence during unscripted speech. This phenomenon is best explained by the cognitive load frameworks described by Song et al. (2023), Parveen (2025), and Du & Salaets (2026), where the real-time retrieval of complex linguistic structures under conversational time constraints severely taxes the working memory of intermediate learners. When students are simultaneously forced to manage lexical selection, phonetic precision, and topical coherence, their internal monitoring systems fail to maintain grammatical accuracy. To minimize this systemic limitation, future iterations of the autonomous model must incorporate dedicated micro-scaffolding prompts and automated corrective feedback loops within the Edpuzzle platform to bridge the gap between structural knowledge and spontaneous usage.

### 4.3. Navigating Logistical and Affective Barriers to Learner Autonomy

The qualitative themes isolated from the participant narratives indicated that the successful execution of autonomous digital learning remains highly dependent on external technical stability and internal psychological resilience. Pervasive technical disruptions, such as local bandwidth fluctuations and browser friction on iOS devices, frequently broke the instructional continuity of the independent study blocks. These infrastructural problems confirm the findings of Zhao & Yang (2024), Zhuang & Liu (2022), and Gómez-Carmona et al. (2023), who warned that the operational viability of innovative mobile educational designs is constrained by the quality of the regional digital ecosystem. A similar technological vulnerability was reported by Wong & Li (2025), and Valdivieso & Cid (2025), who argued that digital readiness gaps and software incompatibilities can disrupt student motivation, highlighting why institutions must guarantee technological equity before deploying self-regulated learning frameworks.

Despite these technical and dialectal friction points, the student cohort demonstrated an encouraging level of analytical awareness and resilience. Faced with rapid native speech rates and unfamiliar British accents, the learners proactively utilized the platform's interactive playback controls to pause, review, and re-examine complex listening inputs. This behavior demonstrates that the authentic challenges embedded within the video materials functioned as a desirable learning difficulty that ultimately fostered deeper meta-cognitive engagement. Additionally, the communication apprehension and mild performance anxiety experienced during voice recording tasks align with the classical affective filter theories (Ma, 2022; Udhayachandran & Sreenivasulu, 2025). However, because the Edpuzzle environment allowed for private, low-stakes multiple takes, students successfully bypassed the fear of immediate negative evaluation that typically limits oral production in traditional, physical classrooms. The combination of technology-mediated authenticity and adaptive multimedia support creates a sustainable, learner-centered environment for high-stakes test preparation.

## 5. CONCLUSION

This study successfully engineered and validated a technology-mediated, self-paced learning model by integrating the interactive capabilities of the Edpuzzle platform with culturally and linguistically authentic video materials for IELTS Speaking preparation. Executed through the rigorous four stages of the 4D development framework, the empirical findings confirmed that the model significantly optimizes learners' pronunciation clarity, lexical resource acquisition, and overall oral fluency. However, the operational field trials also exposed a distinct developmental asymmetry, wherein students' high conceptual understanding of grammatical structures did not seamlessly translate into spontaneous grammatical accuracy during real-time speech production. Ultimately, this research demonstrates that while asynchronous multimedia environments effectively foster learner autonomy and lower affective barriers, targeted pedagogical intervention remains vital to mitigate cognitive overload under unscripted conversational time constraints.

The operational success of this autonomous instructional framework offers vital pragmatic insights for EFL educators, high-stakes test coaches, and curriculum designers operating in digitally mediated environments. Practitioners are strongly encouraged to transition away from rigid, teacher-centric oral drills and instead deploy segmented, authentic multimedia tasks that grant learners meta-cognitive control over their individual learning pace. To bridge the documented grammatical performance gap, instructors must intentionally integrate targeted pop-up prompts and localized micro-scaffolding guides within the digital interface to assist students in converting receptive knowledge into active communicative output. Furthermore, higher education institutions must prioritize investment in optimizing regional digital ecosystems and establishing technological equity to sustain the long-term viability of blended self-regulated learning frameworks.

While this investigation offers valuable insights into autonomous language acquisition, its primary limitation resides in the small and localized participant cohort utilized during the restricted dissemination phase. The qualitative and descriptive nature of the field trial naturally restricts the immediate statistical generalizability of the outcomes across broader, more diverse global EFL demographics. Therefore, future research trajectories should employ large-scale experimental or longitudinal research designs involving diverse student profiles, such as working professionals or advanced English users, to evaluate the long-term sustainability of the platform. Subsequent software engineering and design refinements should also focus on developing automated real-time corrective feedback mechanisms, synchronized interactive subtitles, and an expanded array of specialized speaking topics to enrich global user engagement.

## ACKNOWLEDGMENTS

The authors would like to express sincere gratitude to the Directorate General of Research and Development, Ministry of Higher Education, Science, and Technology, for providing research grant funding under the PDP scheme for the year 2025. This study was supported under Contract Numbers 125/C3/DT.05.00/PL/2025, 7980/LL4/PG/2025, and 34/LPPM/USKM/VI/2025.

## REFERENCES

Ahmad, M. O. (2024). A Deep Dive into Self-Regulated Learning: Reflective Diaries Role and Implementation Strategies. *Communications of the Association for Information Systems*, *54*, 868–888. https://doi.org/10.17705/1CAIS.05432

Akram, H., & Abdelrady, A. H. (2023). Application of ClassPoint tool in reducing EFL learners' test anxiety: an empirical evidence from Saudi Arabia. *Journal of Computers in Education*, *10*(3), 529–547. https://doi.org/10.1007/s40692-023-00265-z

Alam, S. (2025). Effects of Multimedia Inputs on Improving the Grammatical Accuracy of Students' Speaking Skills: An Experimental Study. *Theory and Practice in Language Studies*, *15*(5), 1579–1590. https://doi.org/10.17507/tpls.1505.23

Alamri, W. (2025). Evaluating the benefits and challenges of using authentic materials in EFL context for listening purpose. *Frontiers in Education*, *10*, 1611308. https://doi.org/10.3389/feduc.2025.1611308

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

Blhaj, K. M. S., Dhameria, V., & Muazeib, A. I. M. (2025). The Role of Telemedicine in Digital Transformation Based on Patient Perceptions of Healthcare Accessibility and Efficiency. *International Journal of Nursing Information*, *4*(1), 10–20. https://doi.org/10.58418/ijni.v4i1.131

Bogach, N., Boitsova, E., Chernonog, S., Lamtev, A., Lesnichaya, M., Lezhenin, I., Novopashenny, A., Svechnikov, R., Tsikach, D., Vasiliev, K., Pyshkin, E., & Blake, J. (2021). Speech Processing for Language Learning: A Practical Approach to Computer-Assisted Pronunciation Teaching. *Electronics*, *10*(3), 235. https://doi.org/10.3390/electronics10030235

Coney, N., & Isbell, D. R. (2026). A survey of English language proficiency tests in international student admissions at US research-intensive universities. *Language Testing*, *43*(2), 233–252. https://doi.org/10.1177/02655322251348617

Csapó, B., & Molnár, G. (2019). Online Diagnostic Assessment in Support of Personalized Teaching and Learning: The eDia System. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.01522

Darmawansah, D., Hwang, G.-J., Lin, C.-J., & Febiyani, F. (2025). An artificial intelligence-supported GFCA learning model to enhance L2 students' role-play performance, English speaking and interaction mindset. *Educational Technology Research and Development*, *73*(3), 1451–1479. https://doi.org/10.1007/s11423-025-10453-6

Dmitrenko, N., & Panchenko, V. (2025). Integrating Interactive and Mobile Learning Technologies in Pre-Service EFL Teacher Education. *International Journal of Interactive Mobile Technologies (IJIM)*, *19*(17), 129–150. https://doi.org/10.3991/ijim.v19i17.54477

Du, R., & Salaets, H. (2026). Collaborative learning in interpreting: A collaborative cognitive load theory approach. *Across Languages and Cultures*, *27*(1), 17–36. https://doi.org/10.1556/084.2025.01063

Ferrari, F. (2023). Skills mismatch and change confidence: the impact of training on change recipients' self-efficacy. *European Journal of Training and Development*, *47*(10), 69–90. https://doi.org/10.1108/EJTD-06-2021-0072

Gkintoni, E., Antonopoulou, H., Sortwell, A., & Halkiopoulos, C. (2025). Challenging Cognitive Load Theory: The Role of Educational Neuroscience and Artificial Intelligence in Redefining Learning Efficacy. *Brain Sciences*, *15*(2), 203. https://doi.org/10.3390/brainsci15020203

Gómez-Carmona, O., Buján-Carballal, D., Casado-Mansilla, D., López-de-Ipiña, D., Cano-Benito, J., Cimmino, A., Poveda-Villalón, M., García-Castro, R., Almela-Miralles, J., Apostolidis, D., Drosou, A., Tzovaras, D., Wagner, M., Guadalupe-Rodriguez, M., Salinas, D., Esteller, D., Riera-Rovira, M.,

González, A., Clavijo-Ágreda, J., … Bujalkova, N. (2023). Mind the gap: The AURORAL ecosystem for the digital transformation of smart communities and rural areas. *Technology in Society*, *74*, 102304. https://doi.org/10.1016/j.techsoc.2023.102304

Hu, A., Liu, Q., & Daniel, B. (2025). Digital Technologies in Authentic Assessment in Higher Education: A Systematic Literature Review and Narrative Synthesis. *SAGE Open*, *15*(3). https://doi.org/10.1177/21582440251357198

Hussain, T., Wilany, E., Putri, A., Ria, T. N., & Manuhutu, N. (2025). Technology-Driven Interaction and English Language Learning Outcomes: The Mediating Effects of Engagement and Motivation. *International Journal of Educational Qualitative Quantitative Research*, *4*(2), 1–15. https://doi.org/10.58418/ijeqqr.v4i2.161

Kakitani, J. (2025). Oral task repetition research via videoconferencing. *Research Methods in Applied Linguistics*, *4*(3), 100232. https://doi.org/10.1016/j.rmal.2025.100232

Kharroubi, S., & ElMediouni, A. (2024). Conceptual Review: Cultivating Learner Autonomy Through Self-Directed Learning & Self-Regulated Learning: A Socio-Constructivist Exploration. *International Journal of Language and Literary Studies*, *6*(2), 276–296. https://doi.org/10.36892/ijlls.v6i2.1649

Kolapo, P., Ogunsola, N. O., Komolafe, K., & Omole, D. D. (2025). Envisioning Human–Machine Relationship Towards Mining of the Future: An Overview. *Mining*, *5*(1), 5. https://doi.org/10.3390/mining5010005

Kuluşaklı, E., & Genç, G. (2024). L2 communication apprehension and communicative competence of pre-service English teachers. *Heliyon*, *10*(16), e32587. https://doi.org/10.1016/j.heliyon.2024.e32587

Lestari, M. I., Megawati, F., Agustina, S., & Tappha, Y. (2023). Experiencing EdPuzzle-based Listening Activity: Do the EFL students feel satisfied? *Journal of English Teaching, Literature, and Applied Linguistics*, *7*(2), 32. https://doi.org/10.30587/jetlal.v7i2.5820

Liu, M., Hashim, H., & Sulaiman, N. A. (2026). Longitudinal development of L2 speaking accuracy in AI-enhanced smart class: evidence from an error-annotated learner corpus. *System*, *140*, 104048. https://doi.org/10.1016/j.system.2026.104048

Ma, Y. (2022). The triarchy of L2 learners' emotion, cognition, and language performance: Anxiety, self-efficacy, and speaking skill in lights of the emerging theories in SLA. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.1002492

Mahdi, D. A. (2022). Improving Speaking and Presentation Skills through Interactive Multimedia Environment for Non-Native Speakers of English. *Sage Open*, *12*(1). https://doi.org/10.1177/21582440221079811

Marougkas, A., Troussas, C., Krouska, A., & Sgouropoulou, C. (2023). Virtual Reality in Education: A Review of Learning Theories, Approaches and Methodologies for the Last Decade. *Electronics*, *12*(13), 2832. https://doi.org/10.3390/electronics12132832

Maryam, Z., Rehman, A. U., Nawal, A., Shoaib, M., & Ullah, I. (2025). Examining the relationship between IELTS preparation and English language proficiency skills of international students. *Journal of International Students*, *15*(8), 95–120. https://doi.org/10.32674/rv7pqq16

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

Mekheimer, M. (2025). Technological self-efficacy, motivation, and contextual factors in advanced EFL e-learning: a mixed-methods study of strategy use and satisfaction. *Humanities and Social Sciences Communications*, *12*(1), 677. https://doi.org/10.1057/s41599-025-04947-0

Nieminen, J. H., Bearman, M., & Ajjawi, R. (2023). Designing the digital in authentic assessment: is it fit for purpose? *Assessment & Evaluation in Higher Education*, *48*(4), 529–543. https://doi.org/10.1080/02602938.2022.2089627

Nikolopoulou, K. (2023). Self-Regulated and Mobile-Mediated Learning in Blended Tertiary Education Environments: Student Insights from a Pilot Study. *Sustainability*, *15*(16), 12284. https://doi.org/10.3390/su151612284

Nurdiana, N., Roswati, R., Afrizal, A., Gusmuliana, P., & Dermawan, R. (2025). An Exploration of Self-Directed Learning in Speaking English. *Gurindam: Jurnal Bahasa Dan Sastra*, *4*(2), 80. https://doi.org/10.24014/gjbs.v4i2.34273

Ochoa, X. (2022). Multimodal Systems for Automated Oral Presentation Feedback: A Comparative Analysis. In M. Giannakos, D. Spikol, D. Di Mitri, K. Sharma, X. Ochoa, & R. Hammad (Eds.), *The Multimodal Learning Analytics Handbook* (pp. 53–78). Springer International Publishing. https://doi.org/10.1007/978-3-031-08076-0_3

Octaberlina, L. R., Muslimin, A. I., & Rofiki, I. (2022). An Investigation on the Speaking Constraints and Strategies Used by College Students Studying English as EFL Learners. *International Journal of Learning, Teaching and Educational Research*, *21*(9), 232–249. https://doi.org/10.26803/ijlter.21.9.14

Oshimeje, S., & Flores Barahona, I. D. (2025). Applying Authentic Oral Production to Improve Speaking Fluency Through Project-Based Learning Approach on EFL Students. *Ciencia Latina Revista*

*Científica Multidisciplinar*, *9*(4), 2304–2323. https://doi.org/10.37811/cl_rcm.v9i4.18849

Ostovar-Namaghi, S. A., Morady Moghaddam, M., & Veysmorady, K. (2024). Empowering EFL learners through cognitive apprenticeship: A pathway to success in IELTS speaking proficiency. *Language Teaching Research*, 1–29. https://doi.org/10.1177/13621688241227896

Parveen, N. (2025). Evidence-Based Practices of Cognitive Load Management to Enhance Learning. *Psychological Studies*, *70*(2), 374–386. https://doi.org/10.1007/s12646-025-00841-6

Read, J. (2022). Test Review: The International English Language Testing System (IELTS). *Language Testing*, *39*(4), 679–694. https://doi.org/10.1177/02655322221086211

Setiawati, S., Yuliana, Y. G. S., & Ikhsanudin, I. (2025). Improving Students' Engagement in Listening Activities through Edpuzzle. *Lectura : Jurnal Pendidikan*, *16*(1), 42–53. https://doi.org/10.31849/lectura.v16i1.23344

Song, C., Shin, S.-Y., & Shin, K.-S. (2023). Optimizing Foreign Language Learning in Virtual Reality: A Comprehensive Theoretical Framework Based on Constructivism and Cognitive Load Theory (VR-CCL). *Applied Sciences*, *13*(23), 12557. https://doi.org/10.3390/app132312557

Taye, T., & Mengesha, M. (2024). Identifying and analyzing common English writing challenges among regular undergraduate students. *Heliyon*, *10*(17), e36876. https://doi.org/10.1016/j.heliyon.2024.e36876

Thai, H. B. P., & Nguyen, H. B. (2022). Using authentic materials in teaching speaking to grade 10 students: Teachers' perceptions and practices. *European Journal of Foreign Language Teaching*, *6*(2), 140–158. https://doi.org/10.46827/ejfl.v6i2.4396

Twabu, K. (2025). Enhancing the cognitive load theory and multimedia learning framework with AI insight. *Discover Education*, *4*(1), 160. https://doi.org/10.1007/s44217-025-00592-6

Udhayachandran, R., & Sreenivasulu, Y. (2025). Exploring Communication Anxiety in English Language Learning through Affective Filter Theory: Evidence from Indian College Students. *Journal of Cultural Analysis and Social Change*, *10*(4), 1694–1709. https://doi.org/10.64753/jcasc.v10i4.3069

Valdivieso, C. C., & Cid, R. I. R. (2025). Exploring EFL Instructors' Strategies and challenges for learners Autonomy and Self-regulation in blended learning contexts in Chile. *Ciencia y Reflexión*, *4*(2), 2468–2491. https://doi.org/10.70747/cr.v4i2.419

Wilang, J. D., Kitjaroonchai, N., & Seepho, S. (2026). Reading assistant software and its impact on speaking fluency, grammar accuracy, and narrative quality among EFL learners. *Frontiers in Education*, *11*, 1754473. https://doi.org/10.3389/feduc.2026.1754473

Wong, J. M. S., & Li, K. C. (2025). Technology-Enabled Institutional Readiness for Agile-Blended Learning: A Framework for Educational Innovation. *SN Computer Science*, *6*(3), 279. https://doi.org/10.1007/s42979-025-03822-8

Zahara, N. E. Z. N. E. (2025). Using Edpuzzle interactive video to enhance listening skills. *ELT Forum: Journal of English Language Teaching*, *14*(Special Issue), 334–344.

Zhang, Y., & Dong, C. (2024). Unveiling the Dynamic Mechanisms of Generative AI in English Language Learning: A Hybrid Study Based on fsQCA and System Dynamics. *Behavioral Sciences*, *14*(11), 1015. https://doi.org/10.3390/bs14111015

Zhao, Y., & Yang, Z. (2024). Research on collaborative innovation optimization strategies for digitally enabled higher education ecosystems. *PLOS ONE*, *19*(4), e0302285. https://doi.org/10.1371/journal.pone.0302285

Zhuang, T., & Liu, B. (2022). Sustaining Higher Education Quality by Building an Educational Innovation Ecosystem in China—Policies, Implementations and Effects. *Sustainability*, *14*(13), 7568. https://doi.org/10.3390/su14137568