PASAA Journal
Volume 73, Special Issue 2025/2026, 268–317

CogniWrite: AI-Enhanced Smart Handouts for CEFR-Based Language Instruction and Retention in Higher Education

Diannike Putri[a] and Barlian Kristanto[b]

[a]Universitas An Nasher, Cirebon, West Java, Indonesia

[b]Universitas Harapan Bangsa, Purwokerto, Central Java, Indonesia

Article information	
Abstract	Contemporary higher education faces a digital paradox in which technology abundance coincides with declining cognitive engagement, posing particular challenges for English as a Foreign Language (EFL) instruction. This study examines CogniWrite—an AI-enhanced smart handout system integrating CEFR standards with cognitive activation strategies—among Indonesian EFL learners. Theoretically, CogniWrite operationalizes a dual encoding pathway through three synergistic components: manual writing activating motor-visual encoding via the production effect, reflective scaffolding promoting semantic-elaborative processing aligned with Vygotsky's Zone of Proximal Development, and QR-code-mediated adaptive feedback grounded in retrieval practice principles. A quasi-experimental design compared 100 undergraduate students (experimental n = 50, control n = 50) at B1–B2 level over 14 weeks, with pre-test, post-test, and delayed post-test (four-week washout) assessments using CEFR-aligned rubrics. Both productive skills improved substantially: writing ($d = 0.756$, $p < .001$) and speaking ($d = 0.789$, $p < .001$),

E-ISSN: 2287-0024

	with consistent advantages favoring CogniWrite. Writing retention (d = 0.854) surpassed speaking retention (d = 0.668)—diverging from a priori psycholinguistic predictions and interpretable through the dual encoding mechanism. For EFL practitioners, curriculum designers, and policy makers in resource-constrained contexts, the findings demonstrate that pedagogical and technological sophistication are dissociable design dimensions.
Keywords	CEFR implementation; cognitive activation; AI-enhanced instruction; dual encoding pathway; vocabulary retention; EFL pedagogy
APA citation:	Putri, D. & Kristanto, B. (2025/2026). CogniWrite: AI-enhanced smart handouts for CEFR-based language instruction and retention in higher education [Special Issue]. *PASAA, 73*, 268–317.

1. Introduction

Contemporary higher education faces a fundamental paradox: while students possess unprecedented access to digital resources, their cognitive engagement with learning materials has demonstrably declined. For language teachers, this manifests in familiar student behaviors like photographing slides, recording lectures, or relying on screenshots rather than engaging in active processing. Such passive consumption patterns have been linked to diminished retention, with recent evidence suggesting that smartphone-mediated note-taking may impair rather than support distance video learning outcomes (Hefter, 2024).

Neurological evidence reinforces this concern: manual writing activates broader neural connectivity than typing, engaging sensorimotor processes crucial for memory formation (Van der Weel & Van der Meer, 2024). A recent meta-analysis of 24 studies involving 3,005 college students confirmed that handwritten note-taking produces higher academic achievement than typed notes (Hedges' g = 0.248, $p < .001$), even though typing yields greater note volume (Flanigan et al.,

2024). These advantages reflect a broader principle in memory research: actively producing information generates substantially stronger encoding than passive restudy. Karpicke and Roediger (2008) demonstrated that repeated retrieval practice produced approximately 80% recall after one week, compared to 33–36% for repeated studying, with the production effect extending to vocabulary learning where spoken or written items are recognized more accurately than silently read items (Brown & Roembke, 2024). This benefit appears even more pronounced in second-language contexts, where the cognitive effort of producing L2 material strengthens encoding beyond what passive exposure can achieve (Brown & Roembke, 2024; Terai et al., 2021), differences attributable to depth of processing during encoding (Craik & Lockhart, 1972; Mangen et al., 2022) rather than nostalgic preference.

This challenge becomes especially acute in English as a Foreign Language (EFL) contexts such as Indonesia, where learners face compounded affective and pragmatic difficulties. Indonesian EFL learners in higher education experience significant language anxiety that negatively correlates with communication performance (Putri et al., 2025a), as well as a marked gap between linguistic and interactive competencies that constrains effective performance even when grammatical accuracy is adequate (Putri et al., 2025b). Crucially, Terai et al. (2021) demonstrated that retrieval direction (L2-to-L1 versus L1-to-L2) interacts with vocabulary size, such that lower-proficiency learners benefit more from receptive retrieval while higher-proficiency learners benefit more from productive retrieval, a finding with direct implications for the mixed-proficiency cohorts typical of Indonesian higher education classrooms. Meanwhile, although digital technologies have been widely incorporated into Indonesian higher education, integration has remained largely surface-level, with limited attention to how digital materials are pedagogically scaffolded (Rahmadani, 2024). Implementing CEFR standards in such digital contexts reveals an additional obstacle: the framework presupposes active cognitive engagement (Council of Europe, 2020), an assumption

undermined by the passive consumption that characterizes contemporary learning environments.

Recent advances in AI-enhanced language instruction have demonstrated promising effects on learner engagement and proficiency (Ma et al., 2025; Sun, 2023). Adaptive retrieval-practice systems, in particular, have shown that personalizing item repetition schedules produces measurably better retention than non-adaptive alternatives, with modality choice further modulating outcomes for different learner profiles (Wilschut et al., 2025). Nevertheless, these technologies typically emphasize reactive feedback rather than proactive scaffolding for deep meaning internalization. While scaffolding approaches grounded in Vygotsky's zone of proximal development (ZPD) have proven effective in developing self-efficacy and critical thinking (Allagui, 2024; Sun et al., 2023), practical applications integrating scaffolding with cognitive activation strategies within the CEFR framework remain scarce.

Existing CEFR-based digital instruction studies reveal three interrelated limitations that the present research addresses. First, prior work has concentrated predominantly on assessment alignment and content delivery (Suhan et al., 2024), with little examination of how digital materials can be designed to compel the active cognitive processing that the framework's action-oriented philosophy presupposes. Second, although retrieval practice and production effects are well-documented in cognitive psychology (Brown & Roembke, 2024; Karpicke & Roediger, 2008), their integration into CEFR-aligned instructional design, particularly through manual writing tasks scaffolded by adaptive feedback, has not been systematically investigated. Third, existing adaptive learning systems demonstrating retention benefits (Wilschut et al., 2025) typically require sophisticated infrastructure inaccessible in under-resourced EFL contexts, leaving a critical gap for low-cost solutions that nonetheless preserve the cognitive activation principles these systems embody. The present study addresses these three gaps by examining whether a low-cost, AI-enhanced handout system,

integrating manual writing, reflective scaffolding, and QR-code-mediated adaptive feedback within a CEFR-aligned progression, can produce measurable and durable improvements in productive language skills among Indonesian EFL learners.

To address these converging challenges, this study introduces CogniWrite, an AI-enhanced smart handout system integrating CEFR standards with cognitive activation strategies through three synergistic components. First, manual writing tasks promote deep motor-visual encoding compelled by motor speed constraints (Van der Weel & Van der Meer, 2024), grounded in the production effect whereby actively generating linguistic material yields stronger memory traces than passive reading (Brown & Roembke, 2024). Second, reflective scaffolding facilitates semantic-elaborative processing through guided reflection questions aligned with ZPD principles (Sun et al., 2023), operationalizing retrieval practice within a guided context where effortful but successful retrieval enhances long-term retention (Karpicke & Roediger, 2008). Third, an adaptive feedback system delivered via QR code provides differentiated support across three levels, reinforcement, practice, and critical thinking, calibrated to individual student performance (Bagheri-Nesami et al., 2025), responding directly to Terai et al.'s (2021) finding that optimal retrieval direction shifts with learner proficiency. Learning materials are delivered through AI-enhanced smart handouts that transform conventional CEFR-based presentations into scaffolded fill-in-the-blank tasks requiring active reconstruction rather than passive transcription. By integrating manual writing with reflective scaffolding, the system establishes dual encoding pathways, motor-visual during writing and semantic-elaborative during reflection, strengthening lexical-semantic networks for superior retention.

Employing a quasi-experimental design that compares CogniWrite with conventional methodologies, the present study examines the system's effectiveness in enhancing language proficiency among Indonesian EFL learners, with particular attention to vocabulary retention and productive skill development. The findings address both practical challenges of implementing effective CEFR-

based instruction in digital contexts and theoretical questions concerning how structured communicative frameworks can be enhanced through cognitive processing strategies, yielding insights into differential retention patterns across language modalities with direct implications for EFL pedagogy.

2. Literature Review

2.1 Theoretical Foundations of CogniWrite

CogniWrite's design rests on three converging strands of cognitive and educational research that, taken together, generate testable predictions about how productive language skills should respond to scaffolded manual writing tasks. Rather than treating these strands as isolated literatures, the present study positions them as components of a single dual encoding pathway mechanism, the creation of two distinct yet complementary routes for information processing, that emerges from the integration of manual writing with reflective scaffolding.

The first strand concerns the cognitive consequences of handwriting. Manual writing engages broader sensorimotor neural networks than typing, with motor constraints compelling selective information synthesis rather than verbatim transcription (Mangen et al., 2022; Van der Weel & Van der Meer, 2024). A meta-analysis of 24 studies involving 3,005 college students confirmed that handwritten note-taking yields higher academic achievement than typing (Hedges' $g = 0.248$, $p < .001$), even when typing produces greater note volume (Flanigan et al., 2024). For classroom practice, this body of evidence suggests that handwriting activities, though seemingly traditional, engage deeper cognitive processing than digital note-taking, a principle that becomes especially consequential in the digitally mediated learning environments characteristic of contemporary higher education.

The second strand concerns the production effect and retrieval practice. Karpicke and Roediger (2008) demonstrated that repeated retrieval produced approximately 80% recall after one week, compared with only 33–36% for repeated studying alone. This production benefit extends to vocabulary learning, where

actively generated items are recognized more accurately than passively read items, and appears especially pronounced in second-language contexts where the cognitive effort involved in producing L2 material strengthens encoding beyond what passive exposure can achieve (Brown & Roembke, 2024; Terai et al., 2021). Yet motor-visual encoding alone does not fully explain durable retention; the crucial component lies in its integration with elaborative processing that transforms motor-encoded information into meaningful conceptual understanding.

The third strand concerns depth-of-processing theory (Craik & Lockhart, 1972), which holds that memory traces grow stronger as encoding moves from surface features toward semantic elaboration. Within CogniWrite, reflective scaffolding directs students to engage in semantic-elaborative processing of newly written material, complementing the motor-visual encoding established through manual writing. This integration produces what the present study terms a dual encoding pathway: one route grounded in sensorimotor activation during writing, the other in elaborative reflection on what has just been produced.

Taken together, these three strands generate two priori predictions that guide the present study. First, both productive skills (writing and speaking) should benefit from the dual encoding mechanism, given that vocabulary encoded through handwriting is subsequently retrieved during oral production tasks. Second, and following established psycholinguistic theory reviewed in Section 2.3, speaking would conventionally be expected to show stronger retention than writing, given its multimodal processing characteristics and multiple retrieval pathways (Brown & Roembke, 2024; Meyer, 2023). Whether CogniWrite's instructional design alters this conventional pattern is an empirical question that the present study addresses. For language educators, the theoretical stakes are clear: if scaffolded manual writing can reshape the modality-retention relationship, it offers a principled, low-cost route to durable productive skill development in EFL classrooms.

2.2 Scaffolding Theory in Language Learning Contexts

Building on these cognitive mechanisms, the effectiveness of CogniWrite's reflective scaffolding and adaptive feedback components can be understood through Vygotsky's Zone of Proximal Development (ZPD). This foundational concept in educational psychology emphasizes the importance of graduated support in helping learners achieve capabilities beyond what they can accomplish independently. In language learning contexts, scaffolding functions as a bridge between students' actual and potential abilities, facilitating skill internalization through structured support that gradually diminishes as independence increases (Sun et al., 2023).

In practice, CogniWrite implements scaffolding through two integrated mechanisms. First, reflective scaffolding takes the form of reflection questions that guide students to process materials more deeply, thereby promoting active engagement with language content rather than passive reception. Second, adaptive feedback segmented into three levels—reinforcement, practice, and critical thinking—enables students to access support appropriate to their performance through a QR code-based system. This differentiation reflects ZPD principles, whereby learners at different stages require different types of support (Allagui, 2024). The principle is further reinforced by Terai et al.'s (2021) finding that the optimal direction of retrieval practice (receptive versus productive) shifts with learner proficiency, suggesting that adaptive calibration of cognitive demand is not merely a pedagogical convenience but a theoretically grounded necessity in mixed-proficiency cohorts.

What makes this approach particularly effective is that the self-selection mechanism based on quiz results creates learner agency, wherein students actively participate in the learning process by choosing the support level they need. This approach aligns with findings that scaffolding integrated with reflective activities can deepen learning processes (Sun et al., 2023). The system's adaptability helps explain why scaffolding-based instruction tends to yield

consistent results across diverse student populations, as the support is tailored to individual baseline abilities rather than offered as a one-size-fits-all approach (Shao et al., 2023). For educators working with mixed-ability classes, this adaptability offers a practical solution to the persistent challenge of differentiation, a challenge that is especially acute in Indonesian higher education, where heterogeneous English proficiency within a single cohort is the rule rather than the exception.

2.3 Psycholinguistic Processing and Differential Retention

Perhaps more compellingly for language teachers, psycholinguistic research reveals differential processing mechanisms across language modalities, findings that remain largely unintegrated into instructional design. Speaking requires rapid simultaneous activation of vocabulary, grammar, and pronunciation within narrow temporal windows of approximately 200–300 milliseconds between conversational turns (Krenz et al., 2021; Meyer, 2023). This time pressure appears to enhance memory consolidation through heightened arousal, while the multimodal nature of speaking, combining motor articulation with auditory self-monitoring, establishes multiple retrieval pathways for sustained retention (Brown & Roembke, 2024; Ozker et al., 2024). By contrast, writing relies primarily on visual-motor processes, which prove effective for initial learning but, under conventional instructional conditions, may be less durable for long-term retention (Ihara et al., 2021).

These modality-specific characteristics generate a clear prediction: under standard instructional conditions, speaking should outperform writing in long-term retention because of its multimodal processing and multiple retrieval pathways. Yet this prediction holds only when both modalities are taught through comparable cognitive activation strategies. The present study departs from this baseline by embedding writing within a dual encoding pathway, manual writing combined with reflective scaffolding and retrieval practice through scaffolded fill-in-the-blank tasks (Karpicke & Roediger, 2008), while speaking tasks follow CEFR-aligned communicative formats that emphasize fluency over deliberate encoding.

Whether this asymmetry in instructional design reshapes the conventional modality-retention relationship is the central empirical question the present study addresses. Table 1 juxtaposes the priori psycholinguistic predictions with the retention patterns observed in CogniWrite (reported in full in Section 4), allowing readers to evaluate the extent to which instructional design moderates inherent modality characteristics.

Table 1

Psycholinguistic Processing Predictions and Observed Retention in CogniWrite

Modality	Processing Characteristics	Predicted Retention (a priori)	CogniWrite Results (present study)
Speaking	Multimodal (motor + auditory), time pressure (200–300 ms), multiple retrieval pathways	Higher retention expected	$d = 0.668$*** (medium-large effect)
Writing	Visual-motor processes, slower deliberate encoding, single pathway under conventional instruction	Lower retention expected	$d = 0.854$*** (very large effect)

Note. Effect sizes reported in this column are derived from the present study and are presented in detail in Section 4. ***$p < 0.001$.

The pattern observed diverges from the a priori prediction: writing retention in CogniWrite achieved a very large effect, surpassing speaking retention. This divergence is interpretable through the dual encoding mechanism described in Section 2.1, combined with retrieval practice through scaffolded tasks (Karpicke & Roediger, 2008). The implication for instructional design is consequential: a modality's effectiveness depends not only on its inherent processing characteristics but also on how instructional design leverages or compensates for those characteristics. For language educators, this means writing, long viewed as the more effortful modality to teach for retention, can be repositioned as a high-leverage skill when paired with appropriate cognitive scaffolding.

2.4 Integration with the CEFR Framework

Turning to the role of standardized frameworks, CogniWrite integrates CEFR standards as a learning progression guide facilitating systematic skill development, rather than merely as an assessment tool. This integration manifests in two crucial aspects: first, the selection of learning materials based on CEFR descriptors subsequently transformed into AI-enhanced smart handouts with scaffolded tasks; second, the use of CEFR assessment criteria to measure language proficiency consistently throughout the intervention (Council of Europe, 2001, 2020).

The action-oriented approach central to CEFR philosophy, emphasizing learners' active roles as social agents through real-world tasks (Council of Europe, 2020), aligns closely with cognitive activation strategies in CogniWrite. When students engage in task-based manual writing and reflective activities, they perform meaningful language use consistent with the CEFR principle that language is learned through real communication and purposeful tasks. Yet a persistent obstacle remains: implementing CEFR standards in digital contexts presupposes active cognitive engagement, an assumption increasingly undermined by the passive information consumption that characterizes contemporary learning environments (Council of Europe, 2020; Hefter, 2024). Existing CEFR-based digital instruction has concentrated predominantly on assessment alignment and content delivery (Suhan et al., 2024), with comparatively little attention to how digital materials might be designed to compel the active processing the framework presupposes.

In the present study, vocabulary retention was measured through CEFR-aligned speaking rubrics (Council of Europe, 2020; Fulcher et al., 2011), integrating vocabulary with grammatical accuracy, fluency, and pronunciation. The effect sizes reported in Section 4, large initial gains for both writing and speaking, followed by substantial retention effects, are consistent with the proposition that integrating CEFR standards with cognitive activation strategies can facilitate sustained skill

development. By treating CEFR descriptors as a scaffold for cognitive activation rather than as an assessment endpoint, CogniWrite operationalizes the framework in a way that addresses the active-engagement assumption directly. For practitioners, this reframing suggests that CEFR alignment and cognitive activation are not competing design priorities but complementary ones, provided that instructional materials are designed to compel, rather than merely permit, active processing.

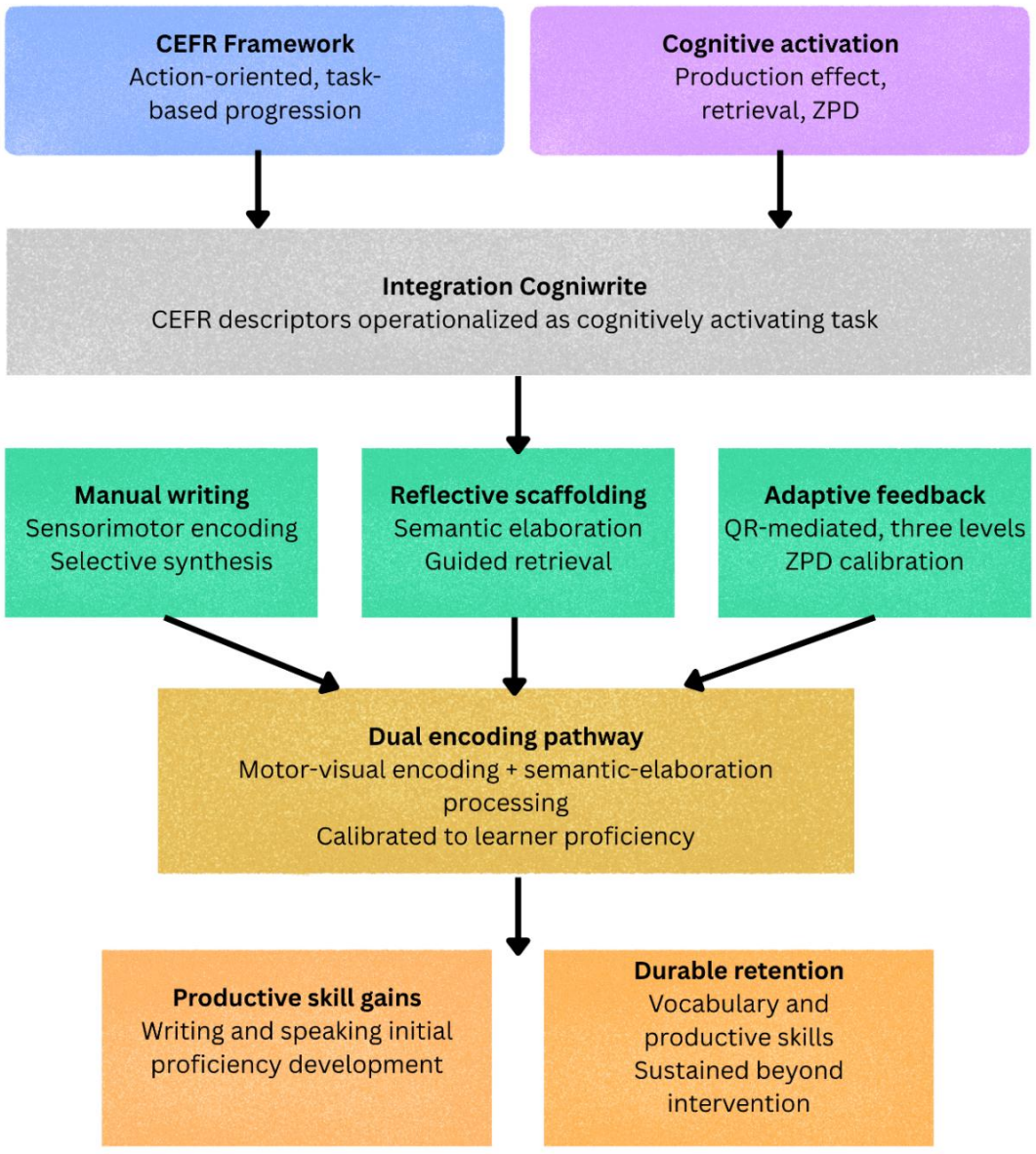
2.5 CogniWrite's Working Model: Theoretical Integration

CogniWrite operates as an integrated system in which three primary components—manual writing, reflective scaffolding, and adaptive feedback—function synergistically to facilitate language learning through distinct yet convergent cognitive activation pathways. Manual writing activates motor-visual encoding (Van der Weel & Van der Meer, 2024); reflective scaffolding facilitates semantic-elaborative processing (Craik & Lockhart, 1972; Sun et al., 2023); and adaptive feedback maintains optimal challenge levels by calibrating cognitive demand to individual learner proficiency (Allagui, 2024; Terai et al., 2021).

The theoretical expectation is that these three components, working in concert, produce robust memory traces by establishing complementary encoding pathways rather than redundant ones. Manual writing alone activates motor-visual encoding but does not guarantee semantic depth; reflective scaffolding alone promotes elaboration but lacks the sensorimotor anchor that strengthens memory traces; adaptive feedback alone maintains engagement but does not in itself drive deep encoding. Their integration is therefore not additive but synergistic, in the sense that each component compensates for the limitations of the others. Figure 1 illustrates how CEFR principles and cognitive activation strategies converge through these three components to produce dual encoding pathways and durable productive skill outcomes.

Figure 1

Conceptual Framework of Cogniwrite: Integration of CEFR Principles and Cognitive Activation Strategies through Three Synergistic Components



From a practical standpoint, this model possesses scalability advantages because it accesses adaptive feedback through simple QR code technology rather than requiring sophisticated infrastructure. This design choice directly addresses the digital paradox foregrounded in Section 1: technology abundance often reduces, rather than enhances, cognitive effort. By using minimal technology to enable maximal cognitive engagement, CogniWrite offers a model that is

particularly suited to under-resourced EFL contexts where sophisticated adaptive learning infrastructure remains inaccessible (Wilschut et al., 2025). For teacher educators and curriculum designers in such contexts, the model suggests that pedagogical sophistication and technological sophistication are dissociable design dimensions.

2.6 Research Gap and Theoretical Positioning

Synthesizing the literature reviewed in the preceding sections, three interrelated gaps emerge that the present study addresses. First, although the cognitive advantages of manual writing (Flanigan et al., 2024; Van der Weel & Van der Meer, 2024) and the production and retrieval-practice effects (Brown & Roembke, 2024; Karpicke & Roediger, 2008) are well-established in cognitive psychology, their systematic integration into CEFR-aligned instructional design has not been examined. Existing CEFR research has prioritized assessment alignment and content delivery (Suhan et al., 2024) over the question of how digital materials might be designed to compel the active processing the framework presupposes (Council of Europe, 2020).

Second, scaffolding research grounded in Vygotsky's ZPD has demonstrated effects on self-efficacy and critical thinking (Allagui, 2024; Sun et al., 2023), and adaptive retrieval-practice systems have shown measurable retention benefits when item repetition is personalized to individual response patterns (Wilschut et al., 2025). Yet practical applications that integrate scaffolding with cognitive activation strategies within the CEFR framework, and that do so in a form accessible to under-resourced classrooms, remain scarce. This gap is particularly consequential for Indonesian EFL contexts, where mixed-proficiency cohorts and surface-level digital integration (Rahmadani, 2024) constrain the effectiveness of imported instructional models.

Third, while psycholinguistic research has clarified the differential processing characteristics of speaking and writing (Brown & Roembke, 2024;

Meyer, 2023; Ozker et al., 2024), the question of whether instructional design can moderate the conventional modality-retention relationship has received limited empirical attention. The interaction of retrieval direction and learner proficiency (Terai et al., 2021) suggests that modality effects are not fixed but conditional on instructional conditions, yet this insight has not been systematically translated into design principles for productive skill development in EFL settings.

The present study addresses these three gaps by examining whether a low-cost, AI-enhanced handout system, integrating manual writing, reflective scaffolding, and QR-code-mediated adaptive feedback within a CEFR-aligned progression, can produce measurable and durable improvements in productive language skills among Indonesian EFL learners. Theoretically, the study contributes to understanding how cognitive activation strategies interact with skill modalities in technology-enhanced language learning. Practically, it offers a scalable model for resource-constrained educational contexts where pedagogical innovation cannot depend on sophisticated technological infrastructure.

3. Methodology

3.1 Research Design

This study employed a quasi-experimental, non-equivalent control group design with pre-test, post-test, and delayed post-test measurements. Full randomization at the individual level was not feasible because the study was conducted in naturalistic educational settings where intact classroom groups must be maintained for institutional and ethical reasons (Gopalan et al., 2020). This design choice prioritizes ecological validity, the extent to which findings transfer to actual instructional conditions, over strict statistical control, a trade-off that is increasingly recognized as appropriate for applied linguistics research targeting EFL learners in higher education contexts. The absence of full randomization carries known implications for the interpretation of findings. Without random assignment, pre-existing group differences in motivation, prior exposure to English, or learning dispositions cannot be fully ruled out as alternative explanations for

observed outcomes, and selection-history or selection-maturation effects may threaten internal validity.

Consequently, the findings reported here should be read as evidence of CogniWrite's effectiveness within the specific conditions of Indonesian EFL higher education classrooms, with generalization to other institutional contexts requiring further empirical verification. To reduce these threats, three mitigation strategies were implemented. First, both groups were drawn from the same institution and academic year, with course schedules, contact hours, and lead instructors held constant. Second, baseline equivalence was verified through institutional CEFR placement testing prior to group assignment, with the experimental and control groups matched on this primary covariate (see Section 3.2). Third, all assessments were scored by trained ESP lecturers external to the research team, blind to group assignment, reducing the risk of expectancy effects influencing outcome measurement.

The study included 100 undergraduate students assigned to two groups: an experimental group (n = 50) receiving the CogniWrite intervention and a control group (n = 50) following conventional instruction. Both groups used materials from the Interchange textbook (Intermediate level). Detailed participant characteristics, recruitment procedures, and inclusion criteria are reported in Section 3.2. The CogniWrite intervention comprised three integrated components. First, manual writing tasks on structured handouts activated motor-visual encoding and selective information processing (Van der Weel & Van der Meer, 2024). Second, reflective scaffolding through guided prompts facilitated semantic-elaborative processing aligned with zone of proximal development principles (Sun et al., 2023). Third, adaptive feedback delivered via QR codes provided tiered support across three levels, reinforcement, practice, and critical thinking, calibrated to individual performance (Bagheri-Nesami et al., 2025). Instructional materials were delivered through AI-enhanced smart handouts that transformed CEFR-aligned presentation

slides into scaffolded fill-in-the-blank tasks, promoting active reconstruction rather than passive transcription.

The experimental timeline comprised 14 weeks of active intervention, followed by mid-term and final examination periods, and then a four-week washout period prior to delayed testing. The four-week interval was selected on converging theoretical and practical grounds. Theoretically, retrieval-practice research indicates that shorter delays risk capturing residual encoding effects rather than consolidated memory, whereas a four-week gap allows the initial forgetting curve to stabilize while remaining short enough to attribute retained gains to the intervention rather than extraneous learning (Karpicke & Roediger, 2008). This is consistent with distributed-practice principles in second-language acquisition, which suggest that retention measured after an extended gap reflects consolidated rather than transient learning gains (Kakitani & Kormos, 2024). Practically, the interval coincided with the institutional post-examination break in the Indonesian academic calendar, during which no formal English instruction occurred in the participating courses. This naturally bounded period ensured that no parallel instructional input could confound retention measurement. Both groups received equivalent instructional time under comparable classroom conditions, with assessments conducted using established CEFR criteria (Council of Europe, 2001, 2018).

3.2 Participants

The study recruited 100 undergraduate students from a private university in Central Java through convenience sampling with intact classroom groups. Participants represented two academic programs: Management (n = 50) and Faculty of Science and Technology (n = 50), providing disciplinary diversity. All participants were native Indonesian speakers learning English as a foreign language.

Three inclusion criteria were established: (1) English proficiency at CEFR levels B1–B2 based on institutional placement testing, ensuring baseline homogeneity; (2) no enrollment in external English courses within the preceding six months, eliminating confounding from supplementary instruction; and (3) minimum 80% attendance throughout the intervention period.

The selection of B1–B2 learners reflects three considerations specific to this study. First, the host institution administers an internal placement test whose results are mapped onto CEFR descriptors, with B1–B2 forming the dominant proficiency cluster at the undergraduate level. Sampling the intact B1–B2 classes therefore captured the most representative proficiency profile of the target population, while levels below B1 (A1–A2) constituted a smaller and pedagogically distinct cluster requiring different instructional pacing. Second, restricting the sample to a single proficiency band controlled for proficiency-related confounds, given that CogniWrite's instructional pacing, the rate at which lecturers deliver explanations alongside the structured handout, must be calibrated to learners' level; mixing widely divergent bands within one study would have introduced pacing as a confounding variable across pre-test, post-test, and delayed-test comparisons. Third, the outcome measures, particularly the oral performance tasks administered prior to mid-term and final-term examinations, presuppose productive competence at the CEFR Independent User threshold (Council of Europe, 2018), at which learners can produce connected discourse and engage in task-based communicative activities with reasonable autonomy. Sampling at B1–B2 ensured that observed outcomes reflected the intervention itself rather than floor effects arising from insufficient productive resources at lower levels.

No attrition occurred during the study. All 100 participants completed the full measurement sequence through delayed testing, yielding a 100% completion rate. The study received ethical approval from the Research and Community Service Institute of the host university (No: B.LPPM-UHB/839/08/2024). Written

informed consent was secured from all participants prior to data collection, and confidentiality was maintained through anonymization.

3.3 Intervention Implementation

CogniWrite was developed in 2023 by the present three-author research team using AI-assisted prompt engineering to transform CEFR-aligned ESP teaching materials, already refined through annual institutional evaluation, into scaffolded fill-in-the-blank handouts, reflective prompts, and tiered adaptive feedback materials.

The CogniWrite intervention was implemented through AI-enhanced smart handouts developed from CEFR-aligned materials in the Interchange textbook (Intermediate level). Development involved AI-assisted transformation of conventional materials into scaffolded fill-in-the-blank tasks, reflective prompts, and structured learning sequences.

Manual writing activities required students to complete structured handouts with designated spaces for active processing, compelling selective information processing through motor constraints. Reflective scaffolding incorporated guided prompts that encouraged students to relate new vocabulary to personal experiences and identify patterns across word forms, establishing dual encoding pathways that strengthened lexical-semantic networks.

The adaptive feedback system utilized QR codes to deliver differentiated support. Students self-selected their support level based on formative quiz results: struggling learners received reinforcement through simplified explanations; intermediate learners engaged with practice exercises; advanced learners received critical thinking extensions. This self-selection mechanism fostered learner agency and self-regulated learning (Shao et al., 2023). Throughout the 14-week intervention, both groups covered identical curricular content; the difference lay in pedagogical delivery method. Each weekly instructional session followed a 90-

minute structure in which lecturer-led explanation of CEFR-aligned content occurred concurrently with student completion of the scaffolded handout, followed by reflective scaffolding and adaptive feedback engagement.

3.4 Instruments and Measurement Procedures

Assessment instruments aligned with the CEFR framework measured productive language skills. The writing rubric assessed four components—Task Achievement, Coherence and Cohesion, Lexical Resource, and Grammatical Range and Accuracy—each contributing 25% to the total score. The speaking rubric assessed four parallel components: Fluency and Coherence, Vocabulary Use (which captured vocabulary retention), Grammatical Range and Accuracy, and Pronunciation, with equal weighting. Vocabulary retention was operationalized through productive use in authentic communicative contexts, in line with the principle that vocabulary mastery is most validly measured through use rather than recognition (Fraser, 1999; Nation, 2001).

3.4.1 Validity of the Instruments

The rubrics were adapted from CEFR descriptors (Council of Europe, 2001, 2018) and the analytic speaking-rubric framework developed by Fulcher et al. (2011), both of which carry well-established content and construct validity from extensive prior use in international language assessment research. Adaptation ensured contextual relevance to Indonesian tertiary ESP settings and produced descriptors practical for routine classroom assessment. Content validity was further established through expert review by two ESP lecturers from another Indonesian university, both active in an Indonesian ESP community of practice. The reviewers evaluated the rubrics on three criteria: descriptor clarity, alignment with CEFR B1–B2 level descriptors, and feasibility of classroom implementation. Their feedback led to minor refinements in descriptor wording but did not alter the underlying CEFR or Fulcher et al. (2011) constructs. The rubrics also reflect iterative refinement through institutional practice: at the host institution, ESP rubrics across eleven study programs are reviewed and revised annually based on

implementation feedback, and the rubrics used in this study had therefore undergone several cycles of practical refinement prior to the present implementation.

3.4.2 Pilot Testing

While a separate formal pilot study was not conducted, both instruments had undergone equivalent prior validation through routine use in ESP assessment with earlier Management cohorts at the host institution. This prior implementation confirmed task clarity, appropriate time allocation, and administrative practicality, with minor adjustments to task wording and timing incorporated before the main study began.

3.4.3 Data Collection Timeline

Data collection followed a structured timeline aligned with the institutional academic calendar. Weeks 1–5 covered the first phase of CogniWrite instruction, followed by speaking assessment in Weeks 6–7 and the mid-term writing examination in Week 8. Weeks 9–13 covered the second phase, followed by speaking assessment in Weeks 14–15 and the final writing examination in Week 16. A four-week washout period preceded delayed testing (see Section 3.1 for justification of the washout interval).

3.4.4 Test Administration

The writing test was administered in a 90-minute session using prompts calibrated to equivalent difficulty levels based on CEFR B1–B2 descriptors. The speaking test was administered individually, lasting 15–20 minutes per student, and followed a standard format comprising a warm-up segment, structured tasks, and free discussion. Administration protocols were applied consistently across all testing phases and both groups to minimize measurement error.

3.4.5 Scoring Procedures

All assessments were scored by trained ESP lecturers external to the research team, using the validated rubrics with rater identity blinded to group assignment. Audio recordings were retained for speaking tasks and anonymized scripts for writing tasks, allowing repeated review where necessary. Inter-rater agreement was monitored through ongoing rater calibration and consistency checks throughout the scoring period, with agreement consistently exceeding .85 across measurement points.

3.5 Data Analysis

Data were analyzed using jamovi (version 2.4) with $\alpha = .05$. Within-group comparisons employed paired samples t-tests to evaluate: (1) pre-test to post-test changes, measuring immediate intervention effects; and (2) post-test to delayed test changes, assessing retention following the washout period. Between-group analyses compared effect sizes at each measurement point.

Effect sizes were calculated using Cohen's d, with $d > 0.65$ established as the threshold for substantial intervention impact. Ninety-five percent confidence intervals accompanied all effect size estimates. Retention analyses employed two complementary metrics: long-term impact (pre-test to delayed test) quantified total learning gains persisting post-intervention, while retention rates (post-test to delayed test) indicated the proportion of immediate gains maintained following the washout period. Differential retention patterns between groups provided evidence regarding CogniWrite's capacity to facilitate genuine language internalization versus transient performance enhancement.

4. Results/Findings

4.1 Immediate Intervention Effects

4.1.1 Within-Group Changes (Pre-test to Post-test)

Paired samples analysis revealed significant improvements in both productive skills within the experimental group following the 14-week intervention.

Writing proficiency demonstrated a mean improvement of 14.27 points (Cohen's d = 0.756, $p < .001$), and speaking proficiency exhibited a gain of 18.83 points (Cohen's d = 0.789, $p < .001$). Both skills exceeded the predetermined Cohen's d > 0.65 threshold for substantial intervention impact. Table 4.1 summarizes these changes.

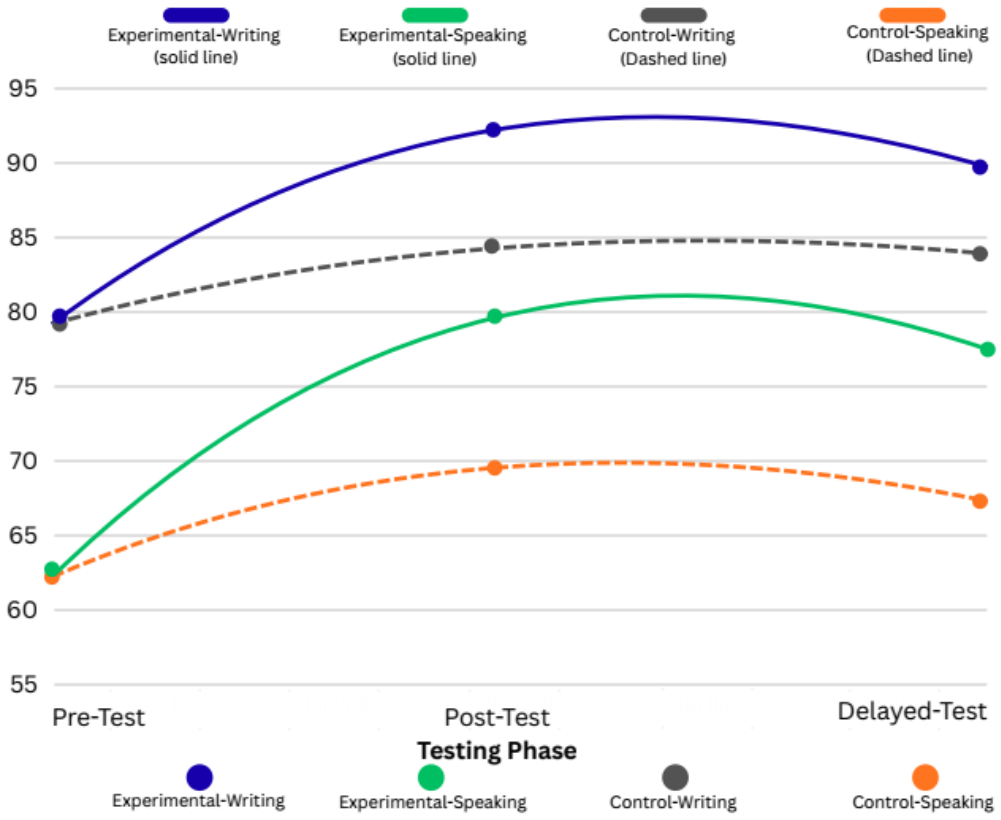
Table 2

Pre-test and Post-test Comparison for Experimental Group

Skills	Mean Pre-test	Mean Post-test	Mean Difference	Cohen's d	p-value
Writing	78.77	93.04	+14.27	0.756	<.001
Speaking	61.73	80.56	+18.83	0.789	<.001

Figure 2

Productive Language Skills Performance Trajectories



4.1.2 Between-Group Comparisons

Between-group comparisons revealed consistent advantages for the CogniWrite intervention across both language skills. The experimental group achieved substantially larger effect sizes than the control group in both writing ($\Delta d = +0.41$) and speaking ($\Delta d = +0.36$). Table 4.2 presents the comparative effect sizes.

Table 3

Effect Size Comparison Between Experimental and Control Groups

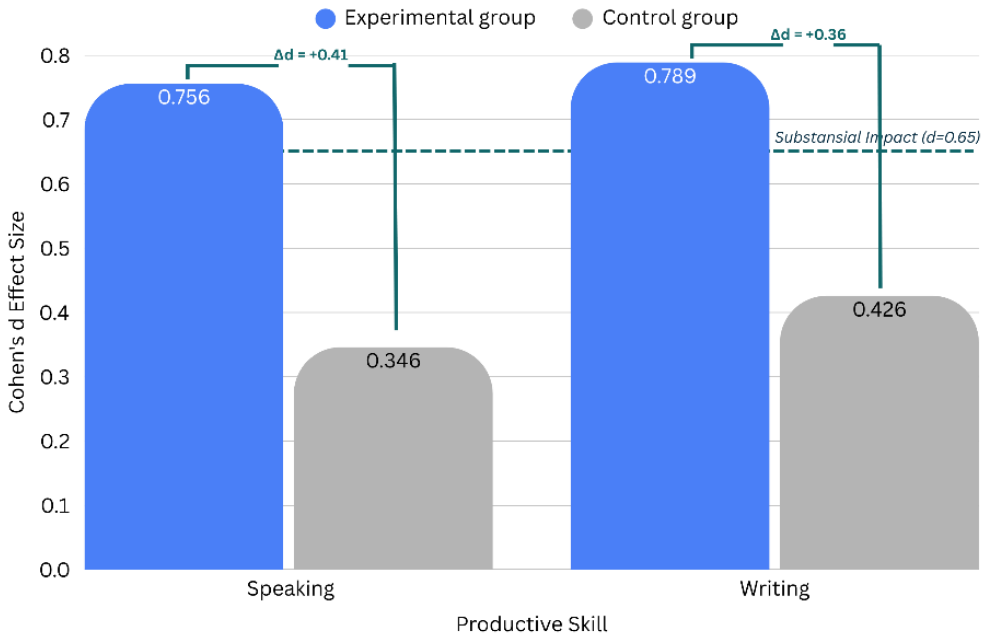
Skills	Experimental (d)	Control (d)	Difference	Statistical Significance
Writing	0.756	0.346	+0.41	$p < .001$
Speaking	0.789	0.426	+0.36	$p < .001$

Note. Both groups received equivalent instructional time (14 weeks, 3 hours/week) and covered identical curricular content.

The control group reached only small effects ($d < 0.5$), while the experimental group reached medium-to-large effects approaching the large threshold ($d = 0.8$). Figure 4.1 shows performance trajectories, and Figure 4.2 visualizes the effect sizes against Cohen's (1988) thresholds.

Figure 3

Cohen's d Effect Sizes: Experimental vs Control Groups



Note Δ d values represent the difference between experimental control effect sizes. Substantial impact threshold (d > 0.65) follows the predetermined criterion for meaningful intervention effects in education contexts. All comparisons were significant at p < .001

4.2 Long-term Retention Effects

Delayed assessment conducted four weeks post-intervention revealed differential retention patterns across the two productive skills (Table 4.3).

Table 4

Long-term Impact and Retention Analysis

Analysis Type	Writing	Speaking	Statistical Significance
Long-term Impact (Pre→Delayed)	d = 0.854	d = 0.668	Both p < .001
Retention (Post→Delayed)	d = 0.260	d = 0.690	Both p < .001

Note. Long-term impact measures total sustained gains from baseline to delayed test. Retention measures maintenance of immediate gains through the washout period.

Long-term impact analysis (pretest to delayed test) demonstrated substantial skill internalization across both domains, with writing ($d = 0.854$, $p < .001$) and speaking ($d = 0.668$, $p < .001$) both maintaining large effects. Retention analysis (posttest to delayed test) revealed that speaking proficiency maintained gains with a medium-to-large effect ($d = 0.690$, $p < .001$), while writing proficiency showed a smaller but positive effect ($d = 0.260$, $p < .001$).

4.3 Vocabulary Retention Analysis

Vocabulary retention was operationalized through the CEFR "Lexical Resource" criterion within speaking assessment rubrics. Between-group retention analysis demonstrated markedly different patterns (Table 4.4).

Table 5

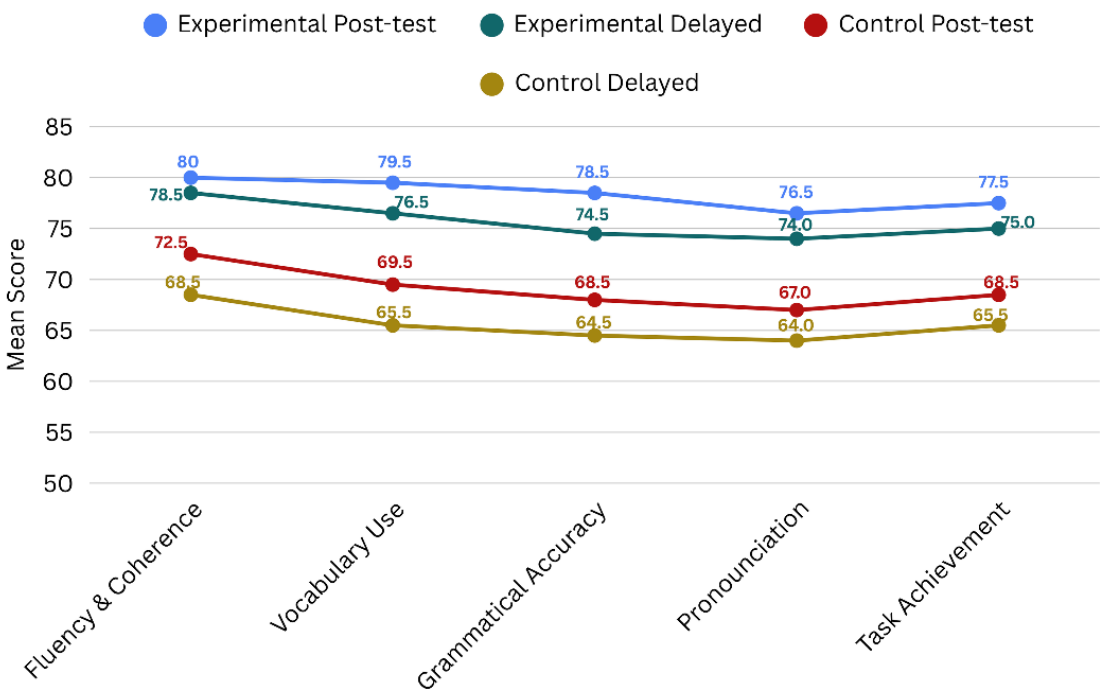
Speaking Proficiency Retention Comparison Between Groups

Group	Speaking Retention (d)	p-value	95% CI	Interpretation
Experimental	0.690	<.001	[12.03, 25.63]	Significant retention
Control	-0.400	.254	[-2.98, 4.56]	Non-significant decline

The experimental group maintained statistically significant gains ($p < .001$) with a confidence interval positioned entirely above zero. In contrast, the control group experienced non-significant decline ($p = .254$), with a confidence interval spanning zero. Figure 4.3 presents vocabulary retention patterns across multiple CEFR descriptors.

Figure 4

Vocabulary Retention Pattern in Speaking Assessment



5. Discussion

5.1 Summary of Key Findings

This study examined the effectiveness of CogniWrite, an AI-enhanced smart handout system integrating CEFR standards with cognitive activation strategies, in enhancing language proficiency among Indonesian EFL learners. The experimental group achieved substantial improvements across both productive skills, with writing ($d = 0.756$, $p < .001$) and speaking ($d = 0.789$, $p < .001$) both exceeding the Cohen's $d > 0.65$ threshold. These gains match earlier studies showing that manual writing and retrieval practice can produce meaningful learning improvements (Flanigan et al., 2024; Karpicke & Roediger, 2008). Between-group comparisons revealed consistent advantages for CogniWrite (writing $\Delta d = +0.41$; speaking $\Delta d = +0.36$), indicating these improvements cannot be attributed solely to practice effects. This result agrees with earlier findings that CEFR-aligned and scaffolded materials lead to stronger learning outcomes than ordinary teaching materials (Council of Europe, 2020; Sun et al., 2023).

Retention analysis following the four-week washout period revealed an unexpected pattern: writing retention ($d = 0.854$, $p < .001$) surpassed speaking retention ($d = 0.668$, $p < .001$). This goes against the common prediction that speaking, because of its multimodal processing and multiple retrieval pathways, should show stronger retention (Brown & Roembke, 2024; Meyer, 2023). By contrast, the control group showed negligible writing retention ($d = 0.260$, $p < .001$) and speaking retention ($d = 0.690$, $p < .001$). This divergence provides compelling evidence that intervention effects extend beyond immediate performance gains to encompass durable skill internalization, which supports the idea that retrieval practice and reflective scaffolding help learners remember material better than passive learning alone (Karpicke & Roediger, 2008; Sun et al., 2023).

5.2 Unexpected Findings: Writing Retention Superiority

That writing retention exceeded speaking retention constitutes an unexpected result requiring careful interpretation. Existing psycholinguistic literature predicts that speaking should demonstrate superior retention due to unique processing characteristics: simultaneous activation of vocabulary, grammar, and pronunciation within narrow temporal windows (200–300 milliseconds), creating time pressure that enhances memory consolidation through heightened arousal (Krenz et al., 2021; Meyer, 2023). Furthermore, speaking involves multimodal processing, combining motor articulation with auditory self-monitoring, creating multiple retrieval pathways supporting sustained retention (Brown & Roembke, 2024; Ozker et al., 2024). Conversely, traditional writing relies primarily on visual-motor processes that, while effective for initial learning, may be less durable for long-term retention (Ihara et al., 2021).

The present findings suggest this represents an intervention-specific effect. CogniWrite creates particularly optimal conditions for writing retention through dual encoding pathways: motor-visual encoding during manual writing activities, and semantic-elaborative encoding during guided reflection. These dual pathways generate more robust memory traces than single-pathway processing. Crucially,

AI-enhanced smart handouts transform CEFR materials into scaffolded fill-in-the-blank tasks requiring active reconstruction, a form of retrieval practice highly effective for long-term retention (Karpicke & Roediger, 2008). The combination of motor-visual encoding, semantic-elaborative encoding, and spaced retrieval creates synergistic conditions for writing retention within this specific intervention context.

Notably, speaking retention, despite showing a lower effect size, still achieved medium-to-large effects ($d = 0.668$, $p < .001$), indicating substantial sustained learning. Controlled speaking tasks within formal instructional contexts tend to encourage focus on fluency and immediate performance rather than deep lexical encoding. This finding demonstrates that modality-specific processing patterns interact dynamically with instructional design: a modality's effectiveness depends not only on its inherent processing characteristics but also on how instructional design leverages those characteristics.

These results have two practical implications for teaching. First, for CEFR-aligned instruction, the findings show that CEFR descriptors are not only tools for assessment; they can also guide active learning when used with scaffolded tasks. This supports earlier studies showing that CEFR materials work better when they ask students to think actively, not just read or listen (Council of Europe, 2020; Suhan et al., 2024). Second, for AI-supported pedagogy, the results show that simple AI tools, such as QR-code feedback, can help students retain skills without expensive technology (Wilschut et al., 2025). This shows that the value of AI in language classrooms comes from how it supports active learning, not from how advanced the technology is (Sun et al., 2023).

5.3 Theoretical Implications: CEFR Integration with Cognitive Activation

The consistency of large effect sizes across both productive skills, combined with robust retention evidence, provides strong support for CogniWrite's theoretical framework. These findings demonstrate that integrating CEFR

standards with cognitive activation strategies can transcend short-term performance gains and facilitate genuine skill internalization.

The reflective scaffolding component proved effective in facilitating semantic-elaborative processing, aligning with zone of proximal development (ZPD) principles whereby scaffolding provides structured support that gradually diminishes as learner autonomy increases (Sun et al., 2023). The adaptive feedback system, providing differentiated support across three levels via QR code technology, creates learner agency and facilitates self-regulated learning calibrated to individual baseline abilities (Shao et al., 2023).

Of particular theoretical significance is integrating CEFR as a learning progression guide rather than merely an assessment tool. The action-oriented approach central to CEFR philosophy, emphasizing learners' active roles as social agents through real tasks (Council of Europe, 2020), aligns naturally with cognitive activation strategies in CogniWrite. When learners engage in task-based manual writing and reflective activities, they perform meaningful language use consistent with CEFR principles that language is learned through real communication and purposeful tasks. This conceptual alignment explains why both productive skills achieved large effect sizes: learners developed communicative competence holistically rather than practicing isolated skills.

5.4 Practical Implications: Scalability and Implementation

An encouraging finding for resource-constrained institutions is that effective educational innovation need not depend on cutting-edge technology. The use of QR codes for adaptive feedback enables intensive cognitive engagement without sophisticated infrastructure, as evidenced by consistent large effect sizes. Successful implementation across diverse populations, from Management to Science and Technology faculties, demonstrates this approach's adaptability with minimal resources. For institutions operating under resource constraints common

throughout Asia, this demonstrates that meaningful pedagogical innovation can occur within existing technological and financial parameters.

CogniWrite's scalability is further facilitated by AI-enhanced smart handouts. Because conventional CEFR materials can be transformed into scaffolded fill-in-the-blank tasks through AI prompting, instructors can implement this approach consistently without significantly increasing preparation burden. AI serves as a force multiplier for instructor effort rather than a replacement for human judgment, an important distinction for sustainable pedagogical innovation. Several concrete implementation recommendations emerge for classroom practitioners. First, differential retention patterns suggest writing skills can be strengthened through strategic combinations of manual processing and structured reflection, while speaking skills benefit from immediate feedback and real-time processing. Practitioners might strategically deploy CogniWrite where its cognitive activation mechanisms align most strongly with skill-specific processing demands. Second, the adaptive feedback mechanism provides a practical solution for differentiated instruction in large classes by allowing learners to self-select support levels based on formative assessments. Third, the manual writing component offers a low-tech solution maintaining cognitive benefits of active processing, particularly relevant where digital devices are unavailable or culturally discouraged during instruction.

Beyond the classroom, the findings also carry implications for two other groups of stakeholders. For EFL curriculum designers, the results suggest that CEFR descriptors should be treated as a guide for active learning tasks, not only as targets for assessment. Curriculum designers can use this approach to build materials that combine CEFR alignment with cognitive activation, especially for mixed-proficiency classes that are common in Indonesian higher education (Council of Europe, 2020; Rahmadani, 2024; Suhan et al., 2024). This means the design process should focus on how each CEFR descriptor can be turned into a task that asks students to think and produce, not just to receive content.

E-ISSN: 2287-0024

For policy makers, the findings show that meaningful improvements in EFL learning do not require large investments in advanced technology. Low-cost tools, such as QR-code feedback and AI-supported handouts, can produce strong learning outcomes when paired with sound pedagogy (Wilschut et al., 2025). This is especially relevant for education policy in resource-limited regions, where digital gaps are common. Policy makers can support adoption through practical and affordable steps, such as making use of free online webinars offered by digital learning platforms, inviting experts to lead intensive workshops at the institutional level, and encouraging follow-up activities within each institution to support sustainability. Such steps can help close the gap between technology-rich and resource-limited classrooms without requiring expensive infrastructure.

6. Limitations, Future Research and Recommendations

Several study limitations warrant acknowledgment. First, the quasi-experimental design with intact classroom groups limits the ability to fully control potential confounding variables, though consistent positive results across diverse populations and 100% completion rate strengthen internal validity. Practitioners should view findings as reflecting genuine intervention impact, though replication in other contexts would further strengthen confidence.

Second, the four-week washout period, while sufficient for identifying genuine internalization versus temporary improvements, is relatively brief compared to longitudinal studies tracking retention over 6-12 months. Future research with extended follow-up periods would clarify whether observed retention patterns persist, strengthen, or attenuate over time. Practitioners should incorporate periodic review and reinforcement activities to ensure sustained mastery.

Third, vocabulary retention was measured through the "Vocabulary Use" component in CEFR-aligned speaking assessment (Fulcher et al., 2011), providing a comprehensive picture of lexical competence within authentic communicative

contexts. The CEFR Companion Volume (Council of Europe, 2020) designates vocabulary range and vocabulary control as components evaluated in speaking performance, providing theoretically valid basis for integrating vocabulary assessment within the CEFR framework. This integrated approach evaluates vocabulary as learners ultimately need to use it, in spontaneous integrated assessment approach offers practical advantages for classroom purposes

Fourth, this study was conducted within a specific Southeast Asian higher education context with EFL learners at CEFR levels B1-B2. Although learner population diversity demonstrates CogniWrite's adaptability across disciplines, generalizability to other geographic contexts, different proficiency levels, or non-tertiary learning contexts requires further validation. Practitioners working with learners outside the B1-B2 range should implement CogniWrite cautiously, monitoring effectiveness and adjusting scaffolding levels as needed.

Fifth, differential retention patterns between writing and speaking must be interpreted cautiously within this study's instructional design context. Writing activities were intensive and direct, while speaking practice may have received indirect benefits from vocabulary encoded through writing activities. Time-on-task allocation and task design optimization for each modality may influence observed retention patterns. Furthermore, controlled speaking tasks may not optimally support deep lexical encoding necessary for long-term retention. Future research with equivalent time-on-task allocation and exploration of spontaneous speaking contexts is needed to clarify whether this retention pattern represents a CogniWrite-specific design effect or a genuine modality-specific pattern. Until such research is conducted, writing retention superiority should be interpreted as conditional upon this study's specific implementation parameters.

Future research directions include: (1) studies with controlled time-on-task allocation for both modalities to clarify whether writing retention superiority represents an intervention design artifact; (2) exploration of alternative speaking

task designs incorporating spontaneous interaction contexts; (3) neuroimaging studies offering insights into how dual encoding pathways activate different neural networks; (4) comparative research contrasting CogniWrite with digital note-taking tools; (5) exploration of implementation at different proficiency levels; (6) longitudinal studies with extended follow-up periods (6-12 months); and (7) cross-linguistic and cross-cultural applications, particularly adaptation to typologically distinct languages with character-based writing systems.

Despite these limitations, robust conclusions emerge for practitioners. Large effect sizes for both skills at posttest, consistent superiority over control conditions, and positive retention effects provide strong evidence that CogniWrite facilitates meaningful and durable learning gains. The intervention's low-resource requirements and demonstrated adaptability across diverse learner populations make it a viable option for evidence-based pedagogical innovation in varied institutional contexts.

7. Conclusion

This study investigated CogniWrite's effectiveness in enhancing language proficiency and vocabulary retention among Indonesian EFL learners. Employing a quasi-experimental design with 100 students from diverse academic backgrounds, the research revealed that CogniWrite generated substantial improvements across both productive skills: writing ($d = 0.756$, $p < .001$) and speaking ($d = 0.789$, $p < .001$), both exceeding the Cohen's $d > 0.65$ threshold indicative of meaningful intervention impact.

Retention analysis following a four-week washout period revealed an unexpected pattern: writing retention demonstrated a very large effect ($d = 0.854$, $p < .001$) superior to speaking retention ($d = 0.668$, $p < .001$). This finding challenges conventional psycholinguistic predictions that speaking skills, given their multimodal processing characteristics, should exhibit greater retention stability. The observed pattern is attributable to dual encoding pathways created

through CogniWrite's integration of manual writing with reflective scaffolding. Motor-visual encoding during handwriting combines synergistically with semantic-elaborative encoding during reflection, with both processes reinforced by scaffolded fill-in-the-blank tasks facilitating active reconstruction rather than passive transcription.

From a theoretical perspective, this research demonstrates that CEFR-based instruction can be significantly strengthened through strategic integration with psycholinguistic principles and cognitive activation strategies. The findings validate an approach that transcends traditional assessment-oriented CEFR applications by operationalizing the framework as a learning progression guide enhanced through systematic cognitive engagement mechanisms. This integration addresses a critical gap in ELT methodology by demonstrating how established frameworks can be enriched through theoretically-grounded cognitive support.

From a practical standpoint, CogniWrite demonstrates that effective educational innovation need not depend on sophisticated technological infrastructure. The intervention's reliance on QR codes for adaptive feedback enables scalable implementation with minimal resources, a particularly significant advantage for resource-constrained contexts common throughout Asia. The consistency of positive results across diverse student populations further supports the approach's robustness and adaptability across varied institutional settings.

Collectively, these findings provide evidence-based guidance for language educators seeking to optimize instruction through strategic integration of established frameworks with accessible cognitive support mechanisms. The research demonstrates that meaningful pedagogical innovation lies not in technological sophistication but in thoughtful alignment of instructional design with cognitive processing principles. For practitioners working in diverse educational contexts, from well-resourced institutions to those facing significant constraints, CogniWrite offers a viable, evidence-based approach to enhancing

productive language skills and ensuring durable learning outcomes. The intervention's demonstrated effectiveness across diverse academic disciplines and its minimal resource requirements position it as a promising option for widespread implementation in varied ELT settings.

8. About the Authors

Diannike Putri is an ESP lecturer and researcher with over 12 years of experience across healthcare, social science, and technology programs in Indonesian higher education. She holds an M.Hum. in English Linguistics from Universitas Padjadjaran and is currently a doctoral candidate in Applied English Linguistics at Unika Atma Jaya Jakarta. Her research interests include meaningful silence in EFL/ESP classrooms, teacher cognition, sociopragmatics, and CEFR-based language instruction.

Barlian Kristanto is a Lektor (Assistant Professor) and ESP & SLA specialist at Universitas Harapan Bangsa, Purwokerto, Indonesia. He holds an M.Hum. from Universitas Sanata Dharma and has received doctoral offers from the University of Nottingham (UK) and the University of Liverpool (UK). A certified Deep Learning facilitator under Kemendikdasmen, he has led 5 competitive research grants, published 12 journal articles and 9 books, and presented at international conferences in South Korea and Thailand.

9. Acknowledgement

The authors would like to express their sincere gratitude to Dr. Ida Dian, whose contributions during the early stages of the research and manuscript preparation laid a valuable foundation for this work. The authors also wish to acknowledge Thanee Glomjai, PhD, RN (Boromarajonani College of Nursing, Phayao, Thailand) for her generous assistance in the statistical verification and final cross-checking of the manuscript. Both contributions are deeply appreciated.

10. Declaration of AI Use

The authors declare that AI tools (ChatGPT and Claude) were used in preparation of the manuscript only, limited to language editing, grammar correction, and clarity refinement of academic English expression. The authors take full responsibility for the content.

11. References

Allagui, B. (2024). A scaffolding intervention to improve self-efficacy in source based argumentative writing. *Frontiers in Psychology*, *15*, Article 1454104. https://doi.org/10.3389/fpsyg.2024.1454104

Bagheri-Nesami, M., Kamali, M., & Mollaei, A. (2025). Comparison of the effect of quick response code-based training and traditional methods on the knowledge, skill, and satisfaction of the nursing students in using medical device. *Smart Learning Environments*, *12*(10). https://doi.org/10.1186/s40561-025-00364-7

Brown, R. M., & Roembke, T. C. (2024). Production benefits on encoding are modulated by language experience: Less experience may help. *Memory & Cognition*, *52*, 926–943. https://doi.org/10.3758/s13421-023-01510-7

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Routledge. https://doi.org/10.4324/9780203771587

Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge University Press.

Council of Europe. (2018). *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume with new descriptors*. Council of Europe Publishing.

Council of Europe. (2020). *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume*. Council of Europe Publishing. https://rm.coe.int/common-european-framework-of-reference-for-languages-learning-teaching/16809ea0d4

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671–684. https://doi.org/10.1016/S0022-5371(72)80001-X

DeKeyser, R. (2007). Introduction: Situating the concept of practice. In R. M. DeKeyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 1–18). Cambridge University Press. https://doi.org/10.1017/CBO9780511667275.002

Flanigan, A. E., Wheeler, J., Colliot, T., Lu, J., & Kiewra, K. A. (2024). Typed versus handwritten lecture notes and college student achievement: A meta-analysis. *Educational Psychology Review*, *36*, Article 78. https://doi.org/10.1007/s10648-024-09914-w

Fraser, C. A. (1999). Lexical processing strategy use and vocabulary learning through reading. *Studies in Second Language Acquisition*, *21*(2), 225–241.

Fulcher, G., Davidson, F., & Kemp, J. (2011). Effective rating scale development for speaking tests: Performance decision trees. *Language Testing*, *28*(1), 5–29.

Gallardo-del-Puerto, F., & Martínez-Adrián, M. (2022). Task modality effects on the production and elaboration of language-related episodes: A study on schoolchildren's interactions in a foreign language. *Languages*, *7*(4), Article 306. https://doi.org/10.3390/languages7040306

Gopalan, M., Rosinger, K., & Ahn, J. B. (2020). Use of quasi-experimental research designs in education research: Growth, promise, and challenges. *Review of Research in Education*, *44*(1), 218–243. https://doi.org/10.3102/0091732X20903302

Hefter, M. H. (2024). Note-taking fosters distance video learning: Smartphones as a risk and intellectual values as a protective factor. *Scientific Reports*, *14*, Article 16962. https://doi.org/10.1038/s41598-024-67898-7

Hyland, K., & Hyland, F. (Eds.). (2019). *Feedback in second language writing: Contexts and issues* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/9781108635547

Ihara, A. S., Nakajima, K., Kake, A., Ishimaru, K., Osugi, K., & Naruse, Y. (2021). Advantage of handwriting over typing on learning words: Evidence from an N400 event-related potential index. *Frontiers in Human Neuroscience*, *15*, Article 679191. https://doi.org/10.3389/fnhum.2021.679191

Kakitani, J., & Kormos, J. (2024). The effect of distributed practice on second language fluency development. *Studies in Second Language Acquisition*, *46*(3), 770–794. https://doi.org/10.1017/S0272263124000251

Karpicke, J. D., & Roediger, H. L. (2008). The critical importance of retrieval for learning. *Science*, *319*(5865), 966–968. https://doi.org/10.1126/science.1152408

Krenz, V., Sommer, T., Alink, A., Antopolskiy, S., Hinrichs, H., & Fischer, R. (2021). Noradrenergic arousal after encoding reverses the course of systems consolidation in humans. *Nature Communications*, *12*, Article 6054. https://doi.org/10.1038/s41467-021-26250-7

Ma, M., Noordin, N., & Razali, A. B. (2025). Improving EFL speaking performance among undergraduate students with an AI-powered mobile app in after-class assignments: An empirical investigation. *Humanities and Social Sciences Communications*, *12*(1), Article 370. https://doi.org/10.1057/s41599-025-04688-0

Mangen, A., Robinet, P., Olivier, G., & Velay, J.-L. (2022). The impact of handwriting on the neural processing of written language: A review of recent neuroimaging research. *Frontiers in Psychology*, *13*, Article 993070.

Meyer, A. S. (2023). Timing in conversation. *Journal of Cognition*, *6*(1), Article 20. https://doi.org/10.5334/joc.268

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.

Ozker, M., Yu, L., Dugan, P., Doyle, W., Friedman, D., Devinsky, O., & Flinker, A. (2024). Speech-induced suppression and vocal feedback sensitivity in human cortex. *eLife*, *13*, Article RP94198. https://doi.org/10.7554/eLife.94198.3

Putri, D., Sumarni, T., Sukmawati, I. D., & Kristanto, B. (2025a). Foreign language anxiety and nursing clinical communication competence in Indonesian pre-elementary nursing students: A mixed methods study. *Journal of Language and Pragmatics Studies*, *4*(1), 1–13. https://doi.org/10.58881/jlps.v4i1.69

Putri, D., Kristanto, B., & Sukmawati, I. D. (2025b). The sociopragmatic competence paradox: Examining the gap between linguistic and interactive skills in EFL academic presentations in science and technology contexts. *Journal of English Teaching and Linguistic Issues (JETLI)*, *4*(1), 44–65. https://doi.org/10.58194/jetli.v4i1

Rahmadani, M. (2024). Depiction of digital literacy practices by English language education lecturers in EFL classrooms. *Indonesian Research Journal in Education (IRJE)*, *8*(2), 662–677. https://doi.org/10.22437/irje.v8i2.31144

Shao, J., Zhang, Q., Li, X., Feng, L., Gao, M., & Pan, H. (2023). Effects of regulated learning scaffolding on regulation strategies and academic performance: A meta-analysis. *Frontiers in Psychology*, *14*, Article 1110086. https://doi.org/10.3389/fpsyg.2023.1110086

Suhan, H., Papageorgiou, S., Davis, L., & Palmer, A. (2024). Mapping the TOEFL ITP® speaking scores to the levels of the Common European Framework of Reference. *Language Testing*, *41*(3), 439–461. https://doi.org/10.1177/02655322231223498

Sun, L., Kangas, M., Ruokamo, H., & Siklander, S. (2023). A systematic literature review on teacher scaffolding in game-based learning in primary education. *Educational Research Review*, *40*, Article 100546. https://doi.org/10.1016/j.edurev.2023.100546

Sun, W. (2023). The impact of automatic speech recognition technology on second language pronunciation and speaking skills of EFL learners: A mixed methods investigation. *Frontiers in Psychology*, *14*, Article 1210187. https://doi.org/10.3389/fpsyg.2023.1210187

Terai, M., & Yamashita, J. (2021). Effects of learning direction in retrieval practice on EFL vocabulary learning. *Studies in Second Language Acquisition*, *43*, 922–946. https://doi.org/10.1017/S0272263121000346

Van der Weel, F. R. R., & Van der Meer, A. L. H. (2024). Handwriting but not typewriting leads to widespread brain connectivity: A high-density EEG study with implications for the classroom. *Frontiers in Psychology*, *14*, Article 1219945. https://doi.org/10.3389/fpsyg.2023.1219945

Wilschut, T., Sense, F., & van Rijn, H. (2025). Modality matters: Evidence for the benefits of speech-based adaptive retrieval practice in learners with dyslexia. *Topics in Cognitive Science*, *17*, 57–72. https://doi.org/10.1111/tops.12769

11. Appendix

COGNIWRITE HANDOUT SAMPLES

FIRST CONDITIONAL SENTENCE

CogniWrite Smart Handout

CEFR Alignment: Level B1

This lesson addresses CEFR B1 descriptor: *"Can express cause and effect relationships using conditional structures with reasonable accuracy in familiar contexts."*

Communicative Goal:

By the end of this lesson, you will be able to discuss future possibilities and their consequences in everyday conversations about risks, decisions, and plans.

Learning Objective

By the end of this lesson, students will be able to use first conditional sentences for _____ conditions and their _____ results.

SECTION 1: Understanding First Conditional

What is First Conditional?

The first conditional is used to talk about _____ situations in the _____.

Example: If it rains tomorrow, I will stay at home.

Complete the meaning: This sentence means the rain is _____ (possible/impossible), and if it happens, staying home is the _____ result.

📝 REFLECTION CORNER 1

Personal Connection:

1. Think about tomorrow. What is one possible thing that might happen? Write a first conditional sentence about it. *Example: If I wake up early, I will...* Your sentence: _____
2. Why do we use first conditional instead of just saying "I will stay home"? What extra information does the "if clause" give us? Your answer: _____

SECTION 2: Structure of First Conditional

First conditional sentence consists of two parts:

1. Hypothetical Condition (If Clause)

- Uses: _____ _____ tense
- Formula: If + _____

2. Probable Result (Main Clause)

- Uses: _____ _____ tense
- Formula: _____ + verb

Complete Structure: If + _____ tense, _____ + verb

📝 REFLECTION CORNER 2

Pattern Recognition:

1. Look at the structure. The if clause uses **present** tense but talks about the **future**. Why do you think English does this? Your idea: _____
2. Complete this comparison:
    - First Conditional: "If I study, I will pass."
    - Simple Future: "I will pass."

What's the difference in meaning? _____

SECTION 3: Tense Review

Simple Present Tense Review

Affirmative Form - Verbal Sentences:

- I, you, we, they + verb (_____ -s/-es)
- He, she, it + verb + _____

Examples:

- I _____ (take) risks.
- She _____ (take) risks.

Negative Form - Verbal Sentences:

- I, you, we, they + _____ + verb
- He, she, it + _____ + verb

Examples:

- I _____ take risks.

- She _____ take risks.

Simple Future Tense Review

Affirmative Form

Formula: Subject + _____ + verb

Examples:

- You _____ discover new opportunities.
- I _____ take the risk.

Contraction: will = _____

Negative Form

Formula: Subject + _____ _____ + verb

Examples:

- You _____ _____ discover opportunities.
- I _____ _____ take the risk.

Contraction: will not = _____

SECTION 4: First Conditional Sentence Types

1. Affirmative Sentences

Structure 1: If clause + comma + main clause

Example: If you _____ a chance, you _____ discover new opportunities.

Structure 2: Main clause + if clause (no comma)

Example: You _____ discover new opportunities if you _____ a chance.

2. Negative Sentences

If clause negative: If you _____ _____ a chance, you _____ _____ discover opportunities.

Main clause negative: If you _____ a chance, you _____ _____ discover opportunities.

Both negative: If you _____ _____ a chance, you _____ _____ discover opportunities.

3. Interrogative Sentences

Question structure: If + present tense, _____ + subject + verb?

Examples:

- If you take a chance, _____ you _____ new opportunities?
- If the opportunity is clear, _____ you _____ the risk?

Note: Only the _____ clause changes to question form. The _____ clause stays the same.

📝 REFLECTION CORNER 3

Metacognitive Check:

1. Which sentence type (affirmative, negative, or interrogative) is most challenging for you? Why? Your answer:

   _____

2. What strategy helps you remember when to use a comma in first conditional? Your strategy:

   _____

SECTION 5: Practice Section A

Complete the sentences with the correct form:

1. If she _____ (try) a new sport, she _____ (have) a lot of fun.
2. If I _____ (take) an action, _____ the result _____ (be) positive?
3. You _____ (feel) achievement if you _____ (challenge) yourself.
4. If Tom _____ _____ (not decide) to start a hobby, he _____ _____ (not have) a chance to discover talent.
5. _____ she _____ (feel) fulfilled if she _____ (quit) her job?

📝 REFLECTION CORNER 4

Real-World Application:

1. Imagine you're giving advice to a friend who is afraid to try something new. Write TWO first conditional sentences giving your advice. Sentence 1: _____ Sentence 2: _____
2. Self-Check: Look at your Practice Section A answers. How confident do you feel? (Circle one)

   Not confident → Scan QR Code Level 1 (Reinforcement)

   Somewhat confident → Scan QR Code Level 2 (Practice)

   Very confident → Scan QR Code Level 3 (Critical Thinking)

SECTION 6: Practice Section B

Write first conditional sentences using these words:

Example: Take / risk / succeed

Answer: If you take the risk, you will succeed.

1. Learn / new skills / improve / career

---

2. Not speak up / not hear / opinion

---

3. Speak / public / become / confident

---

4. Not take / job offer / not earn / higher salary

---

5. Step / outside comfort zone / grow

SECTION 7: Reading Text Application

Taking Risks: A Path to Growth

Some people enjoy taking risks. They like trying new things, meeting new people, or traveling to unknown places. They believe that if they take a risk, they will learn something new. Risk takers often say that life is more exciting when you step out of your comfort zone.

Other people prefer to stay safe. They worry that if they take a risk, they will fail or get hurt. These people feel better when they know what will happen. For them, safety is more important than surprise.

Taking risks can help you grow as a person. If you face your fears, you will become stronger and more confident. Many people who take risks say they learn important lessons, even if things don't go as planned.

Of course, not every risk is a good idea. If you don't think before you act, you will make a mistake. If you plan carefully, you will reduce the danger. That's why it's important to ask questions, do research, and talk to others before making a big decision.

Complete the first conditional sentences from the reading:

1. If they _____ a risk, they _____ learn something new.
2. If they _____ a risk, they _____ fail or get hurt.
3. If you _____ your fears, you _____ become stronger and more confident.
4. If you _____ _____ think before you act, you _____ make a mistake.
5. If you _____ carefully, you _____ reduce the danger.

📝 REFLECTION CORNER 5

Critical Thinking:

1. The text discusses risk-taking. Do you agree that "if you take risks, you will grow as a person"? Write your opinion using at least ONE first conditional sentence to support your view. Your opinion:

   _____

---

2. Personal Connection: Think of a risk you took in the past. What happened? Now write a first conditional sentence about a future risk you're considering. Past risk: _____ Future risk (first conditional):

   _____

ADAPTIVE FEEDBACK SYSTEM

Based on your performance in Practice Sections A & B, scan the appropriate QR code for personalized support:

LEVEL 1: REINFORCEMENT (For students who scored 0-60% or feel "not confident")

[QR CODE 1 HERE]

What you'll get:

- Simplified explanations with visual aids
- More examples with step-by-step breakdown
- Basic comprehension check questions
- Encouragement and tips for understanding the structure

Scan this if: You're still confused about when to use present vs. future tense, or you made more than 3 mistakes in Practice A.

LEVEL 2: PRACTICE (For students who scored 61-80% or feel "somewhat confident")

[QR CODE 2 HERE]

What you'll get:

- Additional practice exercises with varied contexts
- Error analysis: common mistakes and how to fix them
- Application activities requiring productive use
- Tips for vocabulary retention and fluency

Scan this if: You understand the structure but need more practice to feel confident using it naturally.

LEVEL 3: CRITICAL THINKING (For students who scored 81-100% or feel "very confident")

[QR CODE 3 HERE]

What you'll get:

- Advanced application tasks requiring synthesis
- Critical analysis questions about language use
- Creative extension activities (debates, presentations)
- Peer teaching opportunities and leadership roles

Scan this if: You completed all exercises correctly and want to challenge yourself with more complex, real-world applications.

📝 REFLECTION CORNER 6

Metacognitive Reflection:

1. Which QR code level did you choose? Why? Your answer:

   _____

2. After completing the adaptive feedback exercises, do you feel more confident? What helped you most? Your answer:

   _____

SECTION 8: Quick Check

Circle the correct answer:

1. If it (rain / **rains**), I (stay / **will stay**) home.
2. She (**will be** / is) happy if you (**come** / will come) to the party.
3. If we (**don't hurry** / won't hurry), we (miss / **will miss**) the bus.
4. (Do / **Will**) you help me if I (**ask** / will ask) you?
5. If the weather (**is** / will be) nice, we (go / **will go**) to the beach.

SECTION 9: Homework

Write 5 first conditional sentences about your future plans. Use:

- 2 affirmative sentences
- 2 negative sentences
- 1 interrogative sentence

Example: If I study hard, I will pass the exam.

1. _____
2. _____
3. _____
4. _____
5. _____

Homework Reflection

(Answer these after completing your 5 sentences):

1. Which sentence are you most excited about? Why?

---

2. How will you use first conditional in your life this week? Give one specific example.

Important Notes

When to use First Conditional:

- To talk about _____ situations in the future
- When the condition is _____ to happen
- To express _____ and effect relationships

Remember:

- Use _____ tense in the if clause
- Use _____ tense in the main clause
- Use comma when if clause comes _____
- _____ comma when main clause comes first

🎯 Learning Check - Complete This Self-Assessment:

After finishing this handout, I can:

- Explain what first conditional means
- Write affirmative first conditional sentences correctly
- Write negative first conditional sentences correctly
- Write interrogative first conditional sentences correctly
- Use first conditional in real conversations about future plans
- Explain WHY we use present tense in if clause but future in main clause

If you checked fewer than 4 boxes, review the sections you struggled with and access the Level 1 or 2 QR codes for additional support.