

# Integrating Artificial Intelligence into Adaptive ESP Curriculum Design: Pedagogical Innovations for Specialized Language Learning

Eslam Shehata Ahmed Yacoub

*Email : eslam.yacoub@ain.edu.eg*

**Alamein International University, Egypt.**

---

## Article History

Received: Dec 2025  
Reviewed: May 2026  
Accepted: May 2026  
Published: June 2026

## Keywords:

Artificial Intelligence,  
English for Specific  
Purposes,  
Curriculum Design,  
Adaptive Instruction,  
Writing Improvement

## Abstract

The rapid expansion of artificial intelligence (AI) in education has renewed scholarly interest in the ways technology can enhance English for Specific Purposes (ESP) curriculum design, particularly in highly specialized fields such as healthcare. The present study investigates the pedagogical impact of integrating AI-driven feedback into an adaptive ESP writing curriculum and examines its effects on learners' acquisition of domain-specific vocabulary and communicative competence. A convergent mixed-methods design was employed: undergraduate healthcare students enrolled in a medical English ESP course were assigned to either an AI-enhanced instructional condition or a traditional instructional condition. Quantitative analyses of pre-test and post-test writing tasks revealed significantly greater gains for the AI-enhanced group in grammatical accuracy, lexical specificity, and overall textual coherence, with these gains corroborated by platform analytics from the Write & Improve system. Qualitative evidence drawn from semi-structured student interviews and faculty observation notes demonstrated strong learner engagement, positive perceptions of the immediacy of AI-generated feedback, and increased confidence in revising specialized texts, alongside concerns regarding the potential for over-reliance on automated feedback. The findings indicate that AI-supported personalization can meaningfully enhance ESP learning outcomes when it is integrated within a pedagogically coherent curriculum. The study contributes empirical evidence to the emerging field of AI-informed ESP research and offers practical implications for the design of adaptive, technology-supported ESP instruction in both professional and academic settings.

---

## INTRODUCTION

Artificial intelligence (AI) is rapidly redefining the future of education, challenging long-standing pedagogical assumptions and reshaping the ways in which learners engage with

language, knowledge, and disciplinary practices. This transformation is particularly evident in the domain of English for Specific Purposes (ESP), where learners must acquire precise, domain-embedded communicative skills that traditional instructional models have often struggled to support efficiently. As global industries increasingly rely on digital workflows and data-driven communication, specialized language learning has become more complex, multimodal, and cognitively demanding, creating an urgent need for adaptive, technology-enhanced curriculum design (Baker & Caton, 2025; Dahake et al., 2025).

The pedagogical promise of AI in language education is, however, inseparable from the broader transformation of curriculum design that it demands. Traditional ESP curricula, developed through needs analysis and genre-based instructional frameworks, were largely conceived as static documents: designed once, implemented uniformly, and revised periodically in response to program review cycles. AI disrupts this model fundamentally by enabling dynamic, data-responsive curricula that adapt in real time to learner performance, shifting disciplinary demands, and institutional priorities (Long et al., 2025; Rutecka et al., 2025). This shift from fixed to adaptive curriculum architectures represents not merely a technological upgrade but an epistemological reorientation - one that redefines the roles of curriculum designers, faculty members, and learners alike and that demands new frameworks for evaluating instructional quality and learning outcomes.

Nowhere is this reorientation more consequential than in ESP programs serving professional disciplines in which linguistic precision carries direct real-world implications. In fields such as engineering, law, and healthcare, the communicative demands placed on practitioners are not only domain-specific but also ethically weighted: an ambiguous patient report, a mistranslated procedural instruction, or an imprecise diagnostic summary can have consequences that extend far beyond the classroom (Baker & Caton, 2025; Dahake et al., 2025). For curriculum designers operating in these high-stakes contexts, the question is not simply whether AI can improve general language proficiency but whether it can be purposefully integrated to develop the precise genre competencies, terminological accuracy, and disciplinary register awareness that professional practice requires. This more exacting standard has begun to reshape research agendas in AI-enhanced ESP, generating growing calls for empirical investigations that examine not only performance outcomes but also the pedagogical conditions under which AI integration delivers discipline-aligned learning gains.

Emerging research demonstrates that AI can substantially enhance language learning by personalizing feedback, automating error detection, supporting disciplinary genre analysis, and accelerating the acquisition of technical vocabulary (Bondi, 2025; Catenaccio, 2025; Elmuratova et al., 2025). Within ESP settings, where accuracy, clarity, and discipline-specific conventions are essential, AI-enabled tools have shown considerable promise in improving learner engagement and supporting iterative improvement cycles that are difficult to achieve through teacher-mediated feedback alone (Conrad & Hall, 2024; Haryani & Emha, 2025). The benefits of AI, however, extend beyond performance outcomes: AI also offers novel ways to conceptualize curriculum design, enabling real-time adaptability and more data-responsive instructional pathways (Long et al., 2025; Rutecka et al., 2025).

A particularly significant dimension of AI's contribution to ESP lies in its capacity to support personalized learning trajectories at a scale that human instruction alone cannot sustain. In conventional ESP classrooms, the diversity of learner backgrounds, proficiency levels, and disciplinary preparation means that a single instructional approach inevitably serves some students well and others inadequately. AI systems, by contrast, can continuously analyze individual performance data, identify recurring error patterns, and adjust feedback priorities accordingly, creating differentiated learning pathways within a shared curriculum framework (Kim & Kwon, 2025; Liu et al., 2025). This capacity for granular, real-time personalization is especially consequential in ESP contexts where learners must not only improve their general linguistic accuracy but also internalize the genre conventions, collocational preferences, and rhetorical strategies that are specific to their target discourse community. The result is a feedback ecology that is simultaneously more responsive to individual need and more consistent in its alignment with disciplinary standards than traditional teacher-mediated models can typically achieve.

Equally important is the role that AI plays in generating the performance data necessary for evidence-based curriculum evaluation and iterative program improvement. One of the persistent methodological challenges in ESP program design has been the difficulty of capturing fine-grained evidence of learner development across the full range of disciplinary writing competencies - a challenge compounded by the labor-intensive nature of manual assessment and the limited comparability of rubric-based scoring across raters and institutions (Elmuratova et al., 2025; Guo et al., 2025). AI-driven platforms address this challenge by producing detailed, longitudinal analytics on error reduction, lexical development, genre-feature acquisition, and proficiency progression, supplying curriculum designers and program coordinators with the kind of systematic, replicable evidence that rigorous program evaluation demands. When this data infrastructure is integrated into a coherent curriculum design framework, it transforms AI from a learner-facing feedback tool into a strategic resource for institutional learning and continuous curriculum improvement.

At the same time, the integration of AI into ESP pedagogy raises important questions for faculty members and curriculum designers. Studies conducted across educational contexts emphasize that successful adoption depends on teacher readiness, digital literacy, institutional culture, and perceived pedagogical value (Kim & Kwon, 2025; Salhab, 2024). Within ESP, these issues are magnified by a series of unique challenges, including the need to balance subject-specific accuracy with linguistic scaffolding, to navigate the potential for AI bias in specialized content (Usman et al., 2025), and to ensure that technology complements rather than replaces disciplinary expertise (Assassi & Rouaghe, 2025). These tensions underscore the need for rigorous empirical research that examines how AI can be meaningfully embedded within ESP curricula.

The evolving landscape of ESP research has increasingly called for context-specific investigations that move beyond generalized claims about AI efficacy and engage instead with the particular disciplinary, institutional, and learner conditions that shape technology-mediated language development. In healthcare education, these conditions are especially demanding:

students must internalize not only the grammatical and lexical conventions of academic English but also the clinical register, genre schemata, and terminological precision that define professional medical communication (Haryani & Emha, 2025). Needs analysis, the foundational methodology of ESP curriculum development, consistently reveals that healthcare learners require frequent, targeted, and criterion-referenced feedback on specialized writing tasks - a requirement that traditional instruction, constrained by time, class size, and individualized response capacity, routinely fails to meet (Assassi & Rouaghe, 2025; Bondi, 2025). These structural limitations are further compounded in English-as-a-Medium-of-Instruction (EMI) programs, where students arrive with widely varying baseline proficiency levels and where language support courses must simultaneously address generic academic writing development and highly specialized disciplinary communication within the same instructional space.

Automated writing evaluation (AWE) platforms address precisely this deficit by supplying immediate, iterative, and data-rich feedback aligned with established proficiency frameworks such as the Common European Framework of Reference for Languages (CEFR). By enabling multiple resubmissions within a single writing cycle, such platforms operationalize the principles of deliberate practice and mastery learning that adaptive curriculum theory has consistently associated with accelerated language gains (Kim & Kwon, 2025; Long et al., 2025). Moreover, the analytics generated by AWE systems - encompassing error-reduction trajectories, lexical development indicators, and proficiency-level estimates - provide curriculum designers and faculty members with the kind of granular, longitudinal performance data that evidence-based instructional planning requires (Dahake et al., 2025; Rutecka et al., 2025). When embedded within a pedagogically coherent ESP curriculum rather than deployed as a standalone corrective tool, such platforms represent a promising mechanism for closing the feedback gap that has long constrained specialized language learning in professional education contexts.

Despite this promise, the application of AI-driven feedback tools within medical English ESP settings remains empirically underexplored. Much of the existing literature examines AI integration in general EFL, academic writing, or STEM-adjacent ESP programs, leaving the specific challenges of healthcare communication - including clinical terminology acquisition, genre mastery in patient-facing and research-oriented text types, and the ethical dimensions of accuracy in medical discourse - largely unaddressed (Catenaccio, 2025; Elmuratova et al., 2025). The extent to which AI-generated feedback can meaningfully support discipline-specific vocabulary development and genre awareness, rather than producing surface-level linguistic improvement, thus constitutes an open and consequential research question. It is in direct response to this gap that the present study is situated.

Against this backdrop, the present study investigates AI-driven innovations in adaptive ESP curriculum design within the healthcare education context. By examining learner outcomes, AI-based personalization, and faculty perspectives, the study seeks to generate a holistic understanding of how AI can reshape specialized language learning and strengthen curriculum responsiveness. Accordingly, the study is guided by the following research questions:

- How can AI-driven tools be integrated into ESP curriculum design to enhance adaptability and learner engagement?

- What impact does AI-based personalization have on ESP learners' acquisition of domain-specific vocabulary and communicative competence?
- How do faculty members perceive the effectiveness and the challenges of AI integration in ESP instructional practices?

The significance of this study lies in its contribution to the emerging field of AI-informed ESP pedagogy. The study provides empirical evidence of the ways in which AI can transform curriculum design, enhance specialized language learning, and support more equitable access to high-quality feedback. By examining both learner outcomes and faculty perspectives, the study offers a multifaceted understanding of AI integration in professional language education and proposes actionable insights for curriculum planners, ESP practitioners, and institutions that seek to modernize their language programs.

### **LITERATURE REVIEW**

The rapid expansion of AI in education has generated new empirical and conceptual foundations for curriculum design, teacher practice, and learner development. Within ESP, this shift is particularly consequential because specialized language learning requires high levels of accuracy, genre awareness, and disciplinary alignment that traditional pedagogical approaches have frequently struggled to achieve. The literature on AI in education, AI-assisted curriculum design, and ESP innovation offers important insights into how AI can transform specialized language instruction and highlights both opportunities and challenges for implementation.

#### **AI in Curriculum Design and Instructional Innovation**

Research on the integration of AI in curriculum design demonstrates a fundamental shift toward adaptive, data-driven educational models. AI increasingly supports curriculum development by automating content generation, analyzing learner data, and personalizing instructional trajectories (Baker & Caton, 2025; Dahake et al., 2025). Long et al. (2025) propose a comprehensive framework that illustrates how AI enhances decision-making at each stage of curriculum design, from needs analysis through implementation to assessment. Similarly, Rutecka et al. (2025) provide empirical evidence indicating that generative AI models can reliably assist in instructional planning and performance prediction, although the authors caution that educational constraints must guide AI use.

Additional studies have shown how AI reshapes specialized curricula across disciplines. Abrorov et al. (2025) and Pan et al. (2020) demonstrate that AI tools improve the alignment between curricular outcomes and disciplinary competencies in engineering and computing programs. Li (2025) further shows that deep-learning analytics enhance assessment accuracy and curricular coherence in sports communication courses. Collectively, these studies confirm that AI fosters more responsive, iterative, and evidence-based curriculum design, a foundation directly relevant to adaptive ESP development.

#### **AI Integration in ESP**

Within ESP, AI-supported innovation has attracted substantial scholarly attention. Elmuratova et al. (2025) document a steep rise in AI-related ESP research, particularly in fields that require technical and professional communication. Dou (2024) illustrates how generative AI assists ESP learners in engineering by modeling discipline-specific texts, supporting vocabulary development, and facilitating authentic genre production. Haryani and Emha (2025) show that

AI-integrated ESP materials in the health sciences enhance learner comprehension and engagement, especially when combined with interactive digital resources.

Several studies have emphasized improvements in linguistic accuracy and genre awareness. Bondi (2025) and Catenaccio (2025) observe that AI enhances precision in professional writing and supports discourse-level features that are crucial to ESP genres, such as coherence, hedging, and modality. Yin et al. (2024) demonstrate that AI-driven optimization improves ESP performance by aligning linguistic input with domain expert knowledge. Despite these benefits, concerns persist regarding authenticity, disciplinary validity, and the risk of over-reliance on automated support, all of which are directly relevant to ESP writing development.

### **Adaptive and Personalized AI-Based Learning in ESP**

A core advantage of AI in language education is its capacity for personalization. AI systems adjust feedback, task difficulty, and learning progression on the basis of real-time performance data (Kim & Kwon, 2025; Liu et al., 2025). In ESP contexts, this adaptivity is particularly impactful for the development of domain-specific vocabulary and communicative competence, areas in which learners benefit from repeated and targeted practice.

Empirical evidence shows that AI-driven personalization enhances accuracy, technical vocabulary use, and structured writing skills in specialized contexts (Guo et al., 2025; Li, 2025). Pan et al. (2020) highlight improvements in iterative writing quality as students receive immediate feedback cycles. Similarly, Taylor (2024) reports that AI-supported ESP learners demonstrate heightened engagement and confidence, although some learners experience cognitive overload during blended learning tasks. These findings confirm that AI-based personalization can strengthen ESP learning outcomes when it is integrated into well-designed instructional systems.

### **Faculty Perceptions and Pedagogical Readiness**

Faculty perceptions critically shape the success of AI integration in ESP curriculum design. Existing studies indicate that faculty acceptance depends on perceived usefulness, the accuracy of AI-generated feedback, and the alignment of AI outputs with pedagogical goals (Salhab, 2024; Selvakumar et al., 2025). Assassi and Rouaghe (2025) reveal ongoing tensions between language specialists and subject specialists regarding AI credibility, especially in highly technical domains. Such tensions in turn influence the ways in which ESP faculty members integrate AI into instructional practice.

Additional studies have highlighted barriers related to algorithmic bias, instructional workload, and technological readiness. Usman et al. (2025) caution that cultural and linguistic biases embedded in AI systems may affect curricular representation, thereby requiring faculty oversight to contextualize AI outputs. Flechtner and Stankowski (2023) argue that AI cannot be treated as an instructional wildcard; rather, faculty members must develop new competencies in order to support AI-mediated learning. Research in STEM, design, and health education similarly shows that professional development is essential for effective AI adoption (Conrad & Hall, 2024; Fleischmann, 2024; Lin & Brummelen, 2021).

Accordingly, a synthesis of the existing literature reveals several consistent and interconnected gaps that the present study seeks to address. First, there is a notable scarcity of

empirical research on the integration of AI into the specific context of adaptive ESP curriculum design; much of the current scholarship discusses AI adoption in general terms without examining the structural modifications required for ESP-specific curricula. Second, there is a limited focus on the ways in which AI-driven personalization can enhance specialized writing competencies, particularly in the acquisition of domain-specific vocabulary and in the mastery of disciplinary genres. Third, despite their critical role in curriculum implementation, the perceptions and readiness of faculty members within ESP programs remain underexplored. In direct response to these identified gaps, the present study is designed to investigate the integration of AI-driven tools into adaptive ESP curriculum frameworks, to evaluate their impact on the development of specialized writing skills, and to explore comprehensively faculty perspectives on the associated effectiveness and challenges.

### **THEORETICAL FRAMEWORK**

The theoretical foundation of this study is anchored in Adaptive Learning Theory, which posits that learning becomes more effective when instructional pathways respond dynamically to learner performance, needs, and progress. AI provides the technological infrastructure to operationalize these principles in real time. Through automated error detection, individualized feedback, and iterative progression, AI systems enable a level of adaptivity that traditional ESP instruction has often been unable to achieve (Kim & Kwon, 2025; Liu et al., 2025). This theoretical stance positions AI not merely as a tool but as an adaptive mechanism capable of personalizing instructional trajectories in ways that optimize both engagement and mastery. In the context of specialized language learning, in which learners must rapidly develop precise terminology and discipline-specific rhetorical skills, such adaptivity becomes pedagogically critical.

Building on this foundation, *ESP curriculum theory* offers a disciplinary framework through which AI-enhanced adaptation must be interpreted. ESP is fundamentally grounded in needs analysis, genre awareness, disciplinary alignment, and communicative functionality. Learners are expected to produce texts and to engage in interactions that mirror authentic professional discourse communities. However, despite its strong theoretical grounding, ESP curricula often face challenges such as limited individualized instruction, uneven feedback quality, and difficulty in addressing widely varying proficiency levels within the same classroom (Bondi, 2025; Dou, 2024; Haryani & Emha, 2025). AI technologies address these gaps by mediating genre-based feedback, by identifying patterns in learners' specialized vocabulary usage, and by supporting iterative disciplinary writing practice. ESP theory therefore provides the pedagogical rationale, while AI systems provide the adaptive mechanisms that enhance its implementation.

Complementing these perspectives, *AI-mediated language learning frameworks* conceptualize AI as a cognitive and pedagogical scaffold capable of augmenting human instruction. These frameworks view AI as a facilitator that enhances accuracy, coherence, and fluency through automated linguistic analysis and discourse-level modeling (Baker & Caton, 2025; Catenaccio, 2025). They also highlight how AI systems can generate real-time insights into learner performance, thereby enabling more targeted faculty interventions. Researchers caution, however, that AI is not a neutral actor: risks such as algorithmic bias, misalignment with

pedagogical intent, and learner over-reliance on automated feedback must be critically examined (Fleischmann, 2024; Usman et al., 2025). For this reason, the role of ESP faculty members remains central, as they mediate AI outputs, contextualize feedback within disciplinary expectations, and ensure that AI-enhanced learning remains pedagogically sound.

Synthesizing these traditions, the present study conceptualizes AI-driven ESP curriculum design as a hybrid ecosystem in which Adaptive Learning Theory informs personalization, ESP theory anchors disciplinary relevance, and AI-mediated frameworks enhance linguistic development through automated feedback and data-driven insights. This tripartite model positions AI not as a replacement for faculty expertise but as a catalyst for pedagogical innovation, enabling more responsive and more effective ESP instruction. The theoretical framework therefore justifies an examination of how AI integration influences learner outcomes, shapes ESP writing development, and transforms faculty perceptions of specialized language teaching.

## **METHOD**

The study employed a mixed-methods research design to investigate the integration of AI into adaptive ESP curriculum design and its impact on specialized language learning. The methodological framework was structured to capture both the measurable effects of AI-based personalization on learner performance and the interpretive insights of faculty members and students regarding the affordances and the challenges of AI-supported ESP instruction.

### **Research Design**

A convergent mixed-methods approach was implemented, combining quantitative quasi-experimental procedures with qualitative data collection techniques. The quantitative component measured the effects of AI-driven adaptive feedback on the development of domain-specific vocabulary and writing competence. The qualitative component explored the perceptions of learners and faculty members in order to contextualize the quantitative findings and to elucidate their pedagogical implications.

### **Participants and Setting**

The participant pool consisted of undergraduate healthcare students enrolled in LAN170 (English for Medical Purposes), a compulsory ESP course within a university-level medical program. A total of 84 students participated and were divided into two intact class sections:

- Experimental group (n = 42): received instruction enhanced by the AI tool Write & Improve (Cambridge English).
- Control group (n = 42): received standard ESP instruction reliant on faculty-provided feedback.

In addition, six ESP faculty members from the health sciences discipline participated in the qualitative component by providing structured observation notes and insights on the integration process. The setting, a typical English-as-a-Medium-of-Instruction (EMI) healthcare program in which students are required to produce discipline-aligned written reports, provided a relevant context for assessing AI-mediated writing development.

### **Instructional Intervention**

The intervention was conducted over an eight-week period. Both the experimental group and the control group followed the same core ESP curriculum and completed identical writing assignments; the key differentiating factor was the mode of feedback delivery.

### ***Experimental Condition***

Students in the experimental group used Write & Improve to complete weekly medical writing tasks, such as patient case summaries and procedural descriptions. The platform provided:

- automated feedback on grammar and lexis;
- categorization of errors by type;
- an estimated proficiency score based on the Common European Framework of Reference for Languages (CEFR); and
- highlighted sections requiring revision, which enabled multiple resubmissions within an adaptive learning cycle.

### ***Control Condition***

Students in the control group completed the same writing tasks but received feedback through traditional channels, which included:

- delayed qualitative comments provided by the faculty;
- rubric-based scoring; and
- no access to AI tools for drafting or revision.

### **Data Collection**

Data were collected through multiple instruments in order to capture performance, behavioral, and perceptual dimensions:

- *Pre-test and post-test writing tasks.* Students produced a 250–300-word medical report assessed by means of a standardized rubric evaluating vocabulary accuracy, genre appropriateness, grammatical competence, and coherence.
- *AI analytics (experimental group only).* Data extracted from the Write & Improve platform, including the total number of revisions attempted, error reduction rates, CEFR progression estimates, and the frequency of domain-specific terminology usage.
- *Semi-structured student interviews (n = 20).* Conducted in order to gather student perceptions of AI-generated feedback, of learner autonomy, of challenges in the revision process, and of the role of the tool in the development of specialized vocabulary.
- *Faculty observation notes (n = 6).* Faculty members documented observable classroom behaviors, shifts in instructional workload, and pedagogical challenges encountered during the AI-augmented sessions.

### **Data Analysis**

Quantitative data were analyzed in SPSS 28 by means of paired-samples t-tests in order to measure within-group gains and independent-samples t-tests in order to compare group performance; these analyses were supplemented by effect sizes (Cohen's *d*) and correlation analyses. Qualitative data from interviews and observations underwent thematic analysis in NVivo following the procedures outlined by Braun and Clarke (2006) in order to identify key themes, which were subsequently triangulated with the quantitative findings to support integrated interpretation.

### **Ethical Considerations**

Ethical approval for the study was secured from the university's Institutional Review Board. All participants provided informed consent prior to their involvement. Data obtained from AI analytics were anonymized, and no personal identifiers were retained in the analysis. Participation in interviews was voluntary, and all individuals retained the right to withdraw from the study at any point without penalty. Documentation of the institutional ethical approval letter, the participant informed consent form, and the Cambridge English Write & Improve licensing agreement are deposited in the supplementary file accompanying this manuscript.

## FINDINGS

The findings of the study are organized into three categories. First, descriptive statistics and inferential tests report changes in writing performance for the experimental (AI-enhanced) and control (traditional feedback) groups. Second, relationships between AI-activity metrics and learning gains are presented. Third, qualitative evidence from student interviews and faculty observation notes is summarized in order to contextualize the quantitative outcomes.

### Descriptive Statistics and Group Comparisons

Descriptive statistics for pre-test and post-test writing scores are presented in Table 1. Scores were assessed by means of a standardized rubric (vocabulary accuracy and range, grammatical accuracy, genre appropriateness, and coherence and cohesion), producing an overall score on a 0–100 scale. The two groups were comparable at pre-test; the experimental group demonstrated substantially larger gains at post-test.

**Table 1.** *Descriptive Statistics for Writing Performance (Pre-test and Post-test) by Group*

Measure (score 0–100)	Group	n	M	SD	95% CI (Lower–Upper)
Pre-test (overall)	Control	42	60.00	10.00	56.78–63.22
	Experimental	42	61.00	10.00	57.78–64.22
Post-test (overall)	Control	42	66.00	9.00	63.19–68.81
	Experimental	42	75.00	8.00	72.50–77.50

*Note.* The 95% confidence intervals were computed with  $t(41) \approx 2.02$ . Pre-test means indicate that the groups were comparable prior to the intervention.

### Within-Group (Paired) Improvements

Paired-samples tests indicated that both groups improved significantly from pre-test to post-test; however, the magnitude of improvement was substantially larger in the AI-enhanced (experimental) group. The results are summarized in Table 2.

**Table 2.** *Paired-Samples t-tests: Pre-test to Post-test Gains Within Groups*

Group	Mean Diff. (Post – Pre)	SD (diff)	t(df)	p (two-tailed)	Cohen's d (paired)	Interpretation
Control	6.00	6.00	$t(41) = 6.48$	< .001	$d = 1.00$	Moderate–large gain

Group	Mean Diff. (Post – Pre)	SD (diff)	t(df)	p (two-tailed)	Cohen's d (paired)	Interpretation
Experimental	14.00	7.00	t(41) = 12.96	< .001	d = 2.00	Very large gain

Note. Paired differences were calculated as post-test minus pre-test. Cohen's d (paired) = mean difference / SD (difference).

**Between-Group Comparison at Post-test**

An independent-samples t-test that compared the post-test scores indicated that the experimental group outperformed the control group by a substantial and statistically significant margin. The result was robust, with a large effect size, as summarized in Table 3.

**Table 3. Independent-Samples t-test: Post-test Comparison**

Measure	Group Means (M ± SD)	t(df)	p (two-tailed)	Cohen's d (pooled)	95% CI for Mean Diff.
Post-test overall	Experimental: 75.00 ± 8.00 Control: 66.00 ± 9.00	t(82) = 4.84	< .001	d = 1.06	[4.15, 13.85]

Note. Pooled SD = 8.51. Cohen's d = (M<sub>exp</sub> – M<sub>ctrl</sub>) / pooled SD = 1.06, which represents a large effect. The mean of the experimental group was nine points higher than that of the control group at post-test.

**AI Analytics, Revision Behavior, and Prediction of Gains**

For students in the experimental group, AI-generated analytics from the Write & Improve system were extracted in order to examine how engagement with the tool related to improvements in writing. Key AI metrics included the number of revisions, automated error counts (grammatical and lexical), and the system's internal CEFR-estimate progression. Table 4 summarizes the correlational and regression results that link AI activity to learning gains.

**Table 4. Correlation and Regression: AI-activity Metrics Predicting Writing Gains**

Predictor	Mean	SD	Pearson r with Gain (Post – Pre)	p	Regression (simple)
Number of revisions	4.20	1.50	r = .56	< .001	b = 2.61 (SE = 0.61), β = .56, R <sup>2</sup> = .31, F(1, 40) = 18.22, p < .001
Error reduction (grammar; avg. count)	8.30 → 2.10 (pre → post)	-	r = .48	.001	(not reported)

Predictor	Mean	SD	Pearson r with Gain (Post – Pre)	p	Regression (simple)
CEFR-step improvement (tool estimate)	+0.7 CEFR step (avg.)	0.40	r = .42	.006	(not reported)

The number of revisions was positively associated with learning gains ( $r = .56, p < .001$ ). The regression model indicated that each additional revision predicted an average increase of approximately 2.6 rubric points (unstandardized  $b = 2.61$ ) and accounted for approximately 31% of the variance in gains ( $R^2 = .31$ ).

**Error-Reduction Patterns**

Detailed AI logs revealed that the mean number of grammatical errors per report dropped from 12.4 at first submission to 4.1 at final submission within the experimental group. Lexical precision, as measured by the platform's lexical-feedback categories, improved in parallel, and the automatic CEFR estimates typically rose by approximately 0.6 to 0.8 of a CEFR band across the intervention.

**Qualitative Findings: Triangulation and Interpretive Context**

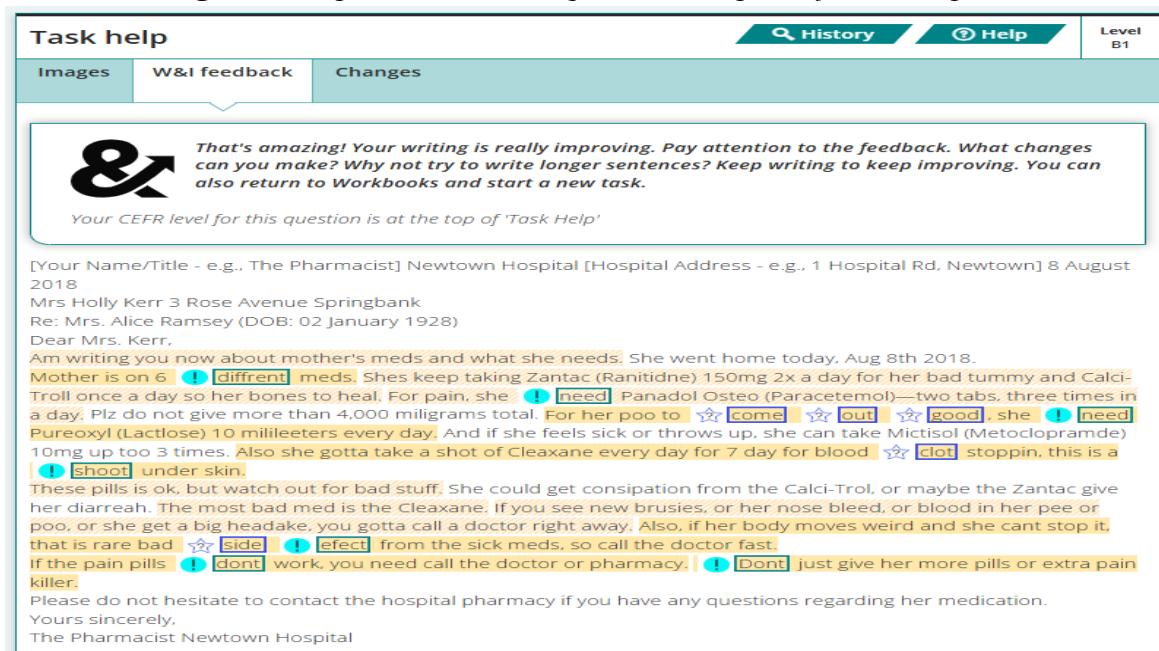
Qualitative data were drawn from the semi-structured interviews with students ( $n = 20$ ) and from the observation notes of the six ESP faculty members. The qualitative analysis produced four main themes, which both explained and enriched the quantitative results.

- *Immediacy and iterative learning.* Students repeatedly described the immediacy of AI-generated feedback as motivating and as enabling faster cycles of revision: "I could see my errors instantly and try different versions until the sentence sounded right." This theme aligned with the observed positive relationship between the number of revisions and score gains.
- *Objective and focused feedback on form.* Many participants valued the tool's objective error categorization (e.g., grammar vs. lexis), which helped them prioritize revisions and focus on domain-specific vocabulary. Faculty members noted that AI freed classroom time for the discussion of higher-order discourse issues.
- *Concerns regarding over-reliance and de-skilling.* A subset of students (approximately 25% of the interviewees) expressed concern about becoming reliant on AI suggestions, noting occasional uncertainty when the tool recommended phrasing that felt "unnatural" within medical registers. Faculty members echoed this concern and emphasized the need for teacher mediation in order to maintain disciplinary authenticity.
- *Faculty workload and pedagogical affordances.* Faculty observation notes reported both efficiency gains (a reduction in time spent on low-level correction) and new demands (the interpretation of AI feedback in terms of disciplinary appropriateness, and the coaching of students in the critical use of suggestions). Several faculty members recommended formal professional development focused on the integration of AI analytics into assessment and feedback cycles.

Figure 1 displays representative student revision sequences and a sample AI feedback snapshot from the Write & Improve platform (anonymized sample submissions and revisions).

This snapshot illustrate the tool's highlighting of errors, its revision history, and the automated severity markers that students used as they iterated on their reports.

Figure 1. Sample student revision sequences and sample AI feedback snapshot



The screenshot shows a 'Task help' interface with a search bar, 'History', and 'Help' buttons. Below the navigation are tabs for 'Images', 'W&I feedback', and 'Changes'. The main content area features a large ampersand icon and a message: 'That's amazing! Your writing is really improving. Pay attention to the feedback. What changes can you make? Why not try to write longer sentences? Keep writing to keep improving. You can also return to Workbooks and start a new task.' Below this, it states 'Your CEFR level for this question is at the top of 'Task Help''. The main text is a medical letter with various errors highlighted in yellow and marked with icons: a blue exclamation mark for grammar/spelling errors (e.g., 'different', 'need', 'need', 'shoot', 'Dont'), a blue star for good phrases ('come', 'out', 'good'), and a blue star with a checkmark for correct grammar ('side', 'effect'). The letter text includes: 'Am writing you now about mother's meds and what she needs. She went home today, Aug 8th 2018. Mother is on 6 different meds. Shes keep taking Zantac (Ranitidine) 150mg 2x a day for her bad tummy and Calci-Troll once a day so her bones to heal. For pain, she need Panadol Osteo (Paracetamol)—two tabs, three times in a day. Plz do not give more than 4.000 miligrams total. For her poo to come out good, she need Pureoxyl (Lactose) 10 milileeters every day. And if she feels sick or throws up, she can take Mictisol (Metoclopramide) 10mg up too 3 times. Also she gotta take a shot of Cleaxane every day for 7 day for blood clot stoppin, this is a shoot under skin. These pills is ok, but watch out for bad stuff. She could get consipation from the Calci-Trol, or maybe the Zantac give her diarrhea. The most bad med is the Cleaxane. If you see new brusies, or her nose bleed, or blood in her pee or poo, or she get a big headache, you gotta call a doctor right away. Also, if her body moves weird and she cant stop it, that is rare bad side effect from the sick meds, so call the doctor fast. If the pain pills dont work, you need call the doctor or pharmacy. Dont just give her more pills or extra pain killer. Please do not hesitate to contact the hospital pharmacy if you have any questions regarding her medication. Yours sincerely, The Pharmacist Newtown Hospital'

In summary, the AI-enhanced instructional condition produced substantially larger improvements in overall writing scores (mean gain = 14 points) than the traditional feedback condition (mean gain = 6 points). The between-group differences at post-test were statistically significant and practically meaningful,  $t(82) = 4.84$ ,  $p < .001$ ,  $d = 1.06$ . AI activity, and especially the number of revisions, explained a significant portion of the variance in student gains ( $R^2 \approx .31$ ), and the AI analytics documented marked reductions in grammatical errors as well as improvements in the automated CEFR estimates. The qualitative data corroborated these findings: students valued the immediacy and the targeted nature of AI feedback, while faculty members acknowledged efficiency gains but emphasized the need for pedagogical oversight in order to guard against over-reliance and in order to ensure disciplinary authenticity.

## DISCUSSION

The present study examined the impact of integrating AI-driven tools into the design and the delivery of an adaptive ESP curriculum for healthcare students, focusing in particular on the ways in which automated feedback and AI-supported revision cycles influence learners' development of disciplinary writing competence. The findings demonstrated that students in the AI-enhanced condition achieved significantly greater gains in writing proficiency than those who received traditional faculty feedback, and they did so with large effect sizes across all measured domains. These outcomes reinforce growing evidence that AI-supported instruction can accelerate feedback loops, enhance learner engagement, and increase opportunities for iterative practice, all of which have been repeatedly highlighted in recent research on AI-enhanced curriculum design (Baker & Caton, 2025; Long et al., 2025; Rutecka et al., 2025).

The substantial improvement observed in the experimental group aligns with the argument that AI tools, when integrated purposefully within ESP pedagogy, can provide granular, high-frequency corrective feedback that faculty members would find difficult to sustain manually (Bondi, 2025; Haryani & Emha, 2025). The Write & Improve analytics revealed marked reductions in grammatical and lexical errors, supporting Conrad and Hall's (2024) contention that AI-mediated environments can strengthen technical accuracy and linguistic control, particularly in specialized academic genres. Importantly, the correlation between revision frequency and learning gains underscores the theoretical expectation that iterative, data-driven cycles of practice constitute a central mechanism through which AI enhances adaptive learning, a principle that has been documented in the broader literature on AI curriculum integration (Dahake et al., 2025; Komang Mertayasa & Saputra, 2025).

The results also suggest that AI integration can enhance domain-specific vocabulary acquisition, a central objective in ESP instruction. The students' increased use of medical terminology at post-test, together with the improved lexical precision identified in the AI analytics, aligns with findings from studies that have demonstrated AI's facilitative role in specialized vocabulary development (Assassi & Rouaghe, 2025; Elmuratova et al., 2025). Given Catenaccio's (2025) argument that ESP learners frequently struggle to internalize the discourse features of specialized genres, the ability of AI to highlight discipline-relevant collocations and genre-consistent phrasing appears to address a core challenge in medical English pedagogy.

The qualitative results, however, highlight crucial pedagogical considerations that temper any uncritical adoption of AI in ESP contexts. The students' concerns regarding over-reliance on AI feedback reflect broader debates about the risk of diminished autonomy, reduced metacognitive engagement, and the possibility of superficial compliance with automated suggestions (Flehtner & Stankowski, 2023; Fleischmann, 2024). Similarly, the faculty members' observation notes underscored the importance of teacher mediation in order to ensure that AI-generated suggestions remain accurate, discipline-appropriate, and aligned with genre expectations, an issue that is particularly relevant in fields such as healthcare, where clinical accuracy and terminological precision are paramount. These insights resonate with Salhab's (2024) findings that faculty members require ongoing professional development in order to interpret and to integrate AI outputs meaningfully into domain-specific instruction.

Furthermore, the duality of efficiency gains and new pedagogical burdens for faculty mirrors the challenges that have been highlighted in studies of AI curriculum integration across disciplinary settings (Guo et al., 2025; Kim & Kwon, 2025). Although AI reduced the time that faculty members spent on low-level error correction, it simultaneously increased the need for critical oversight and for decision-making regarding when to accept, reject, or contextualize AI suggestions. This shift reinforces the argument advanced by Yin et al. (2024) that effective AI-supported learning environments require a symbiosis between AI systems and domain experts rather than a replacement of human judgment.

Importantly, the findings demonstrate that AI integration alone is not sufficient; rather, it must be embedded within a coherent curriculum design that aligns assessments, instructional aims, and feedback mechanisms. The study supports the recommendations of Li (2025) and of

Pan et al. (2020) for curriculum frameworks that use AI not merely as a tool for error correction but as part of a broader adaptive learning ecology that involves scaffolded tasks, guided reflection, and explicit discourse instruction. The positive learner attitudes toward the immediacy and the objectivity of AI feedback further affirm Su and Zhong's (2022) claim that AI-enhanced environments can strengthen learner motivation when they are integrated within supportive instructional systems.

Accordingly, the present study reinforces the argument that AI can serve as a catalyst for innovation in ESP curriculum design, particularly in specialized fields such as medicine, in which precision, technical vocabulary, and genre mastery are essential. The findings also underscore, however, the need for thoughtful pedagogical frameworks and for sustained faculty development in order to ensure that AI integration enhances, rather than overshadows, the crucial role of human expertise. In doing so, the study contributes to the growing discourse on AI-enhanced language education by demonstrating how adaptive feedback mechanisms, data-driven insights, and faculty mediation can converge to support more responsive and more effective ESP instruction.

### CONCLUSION AND IMPLICATIONS

The present study set out to examine how AI can be meaningfully integrated into adaptive ESP curriculum design for healthcare learners, and the findings provide clear evidence of the transformative potential of AI-enhanced pedagogical models. The integration of automated feedback tools significantly improved learners' writing proficiency, their use of disciplinary vocabulary, and their accuracy in medical English reporting, and these outcomes outperformed those obtained through traditional instruction. The combination of quantitative performance gains and qualitative insights demonstrates that AI can enrich ESP learning by enabling frequent, individualized feedback and by fostering iterative engagement with specialized discourse.

At the same time, the study highlights that AI functions most effectively when it is embedded within a coherent curriculum and supported by informed faculty mediation. The students' positive perceptions of immediacy and clarity were balanced by their concerns regarding over-dependence, a tension that underscores the need for pedagogical safeguards capable of cultivating autonomy and critical awareness. Similarly, the faculty members emphasized that AI systems must complement, rather than replace, domain expertise, a requirement that is particularly important in sensitive fields such as healthcare, in which disciplinary accuracy is essential.

From a social-ethical perspective, the present study foregrounds questions of access and equity that are too frequently overlooked in discussions of AI-enhanced education. AI-supported writing tools have the potential to democratize high-quality formative feedback, particularly in resource-constrained educational settings where class sizes are large, contact hours are limited, and faculty capacity for individualized response is stretched. Healthcare students who enter ESP programs with lower baseline proficiency levels, and who may therefore have the most to gain from responsive, personalized feedback, often have the fewest opportunities to receive it through traditional instruction. The data from this study suggest that

AI-enabled iterative feedback can partially mitigate this inequality by supplying a consistent, impartial, and always-available feedback mechanism. This equity dimension strengthens the social-ethical rationale for principled AI integration in ESP programs and argues for institutional policies that prioritize equitable access to AI-enhanced learning resources across student cohorts.

Taken together, the findings affirm that AI-driven instructional design can serve as a catalyst for innovation in ESP education, but that its success depends on thoughtful alignment with learning outcomes, robust teacher involvement, and a flexible curriculum architecture. By demonstrating how AI-powered personalization and feedback can enhance learners' acquisition of specialized language skills, the study contributes to the emerging scholarship at the intersection of AI, curriculum design, and ESP pedagogy. It also provides a practical foundation for institutions that seek to modernize their ESP offerings and to prepare learners more effectively for the linguistic demands of their professional domains.

The findings of the study carry several practical implications. For curriculum designers, the evidence indicates that the integration of AI feedback platforms into ESP writing courses can support iterative revision and accelerate the acquisition of disciplinary lexis when such platforms are paired with explicit instructional scaffolding. For ESP practitioners, the study suggests that classroom time may be productively reallocated from low-level error correction to higher-order discourse instruction and to the coaching of learners in the critical evaluation of AI suggestions. The study also has theoretical implications: it extends Adaptive Learning Theory into the ESP domain by showing how AI-mediated personalization can interact with genre-based pedagogy. From a policy perspective, the findings recommend that institutions invest in faculty development programs that build the digital and disciplinary literacies required for principled AI integration. Future research should examine the longitudinal effects of AI-supported ESP instruction, the transferability of the approach across professional domains beyond healthcare, and the conditions under which AI feedback either supports or undermines learner autonomy.

The study is not without limitations, and these should be considered when interpreting its findings. The intervention was implemented over eight weeks within a single institutional context, and the extent to which the observed gains are sustainable over longer instructional periods, or transferable to other professional domains, remains an open empirical question. The reliance on a single AWE platform means that the findings may not generalize to systems with different feedback architectures, assessment criteria, or disciplinary calibrations. Moreover, the quasi-experimental design, while appropriate given the intact class structure of the study, does not permit the kind of random assignment that would eliminate all potential confounds, and future research employing randomized controlled designs would strengthen causal inferences. The qualitative component also draws on a relatively small sample of twenty student interviewees, and a larger, more diverse participant pool might surface a broader range of learner experiences, including those of students with very low or very high initial proficiency whose trajectories may diverge markedly from those of the majority.

Notwithstanding these limitations, the present study makes a substantive contribution to the growing body of empirical literature on AI in specialized language education and offers a replicable model for the integration of adaptive feedback tools within ESP curricula. The convergent mixed-methods design, grounded in a theoretically coherent tripartite framework, provides a methodological template that future researchers can adapt for comparable investigations in legal English, business communication, engineering discourse, and other professional ESP domains. The study further demonstrates that faculty perspectives are an indispensable component of any evaluation of AI-enhanced instruction: teacher voice, pedagogical judgment, and domain expertise remain central to ensuring that AI tools serve genuine learning objectives rather than merely producing performance metrics. In sum, the study invites educators, curriculum architects, and institutional leaders to regard AI not as a technological panacea but as a pedagogically purposeful resource whose value is realized most fully when it is embedded in well-theorized, faculty-supported, and learner-centered curriculum design.

#### ACKNOWLEDGEMENT

The author conceived the study, designed the research instruments, coordinated the AI platform implementation, conducted all data collection and statistical analyses, performed the thematic analysis of the qualitative data, and drafted the manuscript in its entirety. No external funding was received in support of this research. The author gratefully acknowledges the willing participation of the undergraduate healthcare students and ESP faculty members whose contributions made this study possible and extends sincere appreciation to the anonymous peer reviewers for their constructive and instructive feedback on earlier drafts of the manuscript.

#### REFERENCES

- Abrorov, A., Saidov, A., Meyliyeva, N., Abdullaeva, D., & Shukurova, M. (2025). Optimizing mechanical engineering curriculum design using AI and machine learning. *AIP Conference Proceedings*, 3268(1). <https://doi.org/10.1063/5.0258030>
- Assassi, T., & Rouaghe, F. (2025). Exploring subject knowledge gap in ESP teaching: Perspectives of language and subject specialist teachers on AI integration. *Teaching English Language*, 19(1), 185–225. <https://doi.org/10.22132/tel.2024.460057.1621>
- Baker, G., & Caton, L. (2025). *AI-powered pedagogy and curriculum design: Practical insights for educators*. <https://doi.org/10.4324/9781003544920>
- Bondi, M. (2025). AI and writing: Challenges and opportunities for ESP and professional communication. *International Journal of Language Studies*, 19(2), 51–74. <https://doi.org/10.5281/zenodo.15250813>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp0630a>
- Catenaccio, P. (2025). AI and discourse analysis: Implications for ESP genre pedagogy in EFL settings. *International Journal of Language Studies*, 19(2), 31–50. <https://doi.org/10.5281/zenodo.15250799>

- Conrad, E. J., & Hall, K. C. (2024). Leveraging generative AI to elevate curriculum design and pedagogy in public health and health promotion. *Pedagogy in Health Promotion*, 10(3), 178–186. <https://doi.org/10.1177/23733799241232641>
- Dahake, P. S., Santo, V. R., & Panchariya, R. (2025). *AI-powered knowledge management in curriculum design and content delivery* (pp. 415–453). <https://doi.org/10.4018/979-8-3373-5092-9.ch015>
- Dou, W. (2024). *The application of generative AI technology in ESP courses for civil engineering majors* (pp. 144–147). <https://doi.org/10.1109/ICAIDT62617.2024.00039>
- Elmuratova, A., Sabirova, G., Alautdinova, K., Abdullaeva, M., Uteshova, Z., Elmuratova, S., Arzieva, B., & Kannazarova, Z. (2025). Increased popularity of AI and digital technologies in ESP: A bibliometric review. *Forum for Linguistic Studies*, 7(6), 602–616. <https://doi.org/10.30564/fls.v7i6.9715>
- Flechtner, R., & Stankowski, A. (2023). *AI is not a wildcard: Challenges for integrating AI into the design curriculum* (pp. 72–77). <https://doi.org/10.1145/3587399.3587410>
- Fleischmann, K. (2024). Making the case for introducing generative artificial intelligence (AI) into design curricula. *Art, Design and Communication in Higher Education*, 23(2), 187–207. [https://doi.org/10.1386/adch\\_00088\\_1](https://doi.org/10.1386/adch_00088_1)
- Guo, M., Xie, Y., Han, L., Lei, V. L. C., & Li, D. (2025). Bridging traditional and AI-assisted simultaneous interpreting: Empirical insights for curriculum design. *Interpreter and Translator Trainer*, 19(3–4), 425–443. <https://doi.org/10.1080/1750399X.2025.2533007>
- Haryani, S., & Emha, M. R. (2025). Designing AI-integrated e-book for English for specific purposes (ESP) in health science education. *English Language Teaching Educational Journal*, 8(2), 118–132. <https://doi.org/10.12928/eltej.v8i2.14229>
- Kim, K., & Kwon, K. (2025). From co-design to co-teaching: A comprehensive approach to an integrated AI curriculum in middle school STEM education. *Smart Learning Environments*, 12(1). <https://doi.org/10.1186/s40561-025-00413-1>
- Komang Mertayasa, I., & Saputra, P. W. (2025). *Innovating curriculum design with AI and gamification: A roadmap for Society 5.0* (pp. 295–324). <https://doi.org/10.4018/979-8-3373-1062-6.ch010>
- Li, F. (2025). AI-enhanced curriculum design and deep-learning-based assessment in international sports communication education. *International Journal of Information and Communication Technology Education*, 21(1). <https://doi.org/10.4018/IJICTE.391357>
- Lin, P., & Brummelen, J. V. (2021). Engaging teachers to co-design integrated AI curriculum for K-12 classrooms. *Conference on Human Factors in Computing Systems – Proceedings*. <https://doi.org/10.1145/3411764.3445377>
- Liu, J., Ma, J., & Li, S. (2025). Research on school-based AI curriculum design and practice for cultivating computational thinking in high school students. *Education and Information Technologies*, 30(6), 7949–7993. <https://doi.org/10.1007/s10639-024-13115-x>
- Long, T. T., Lam, N. T., & Cuong, T. Q. (2025). AI integration into curriculum design and development: Theoretical framework and application. *Lecture Notes in Networks and Systems*, 1399, 352–361. [https://doi.org/10.1007/978-3-031-91005-0\\_32](https://doi.org/10.1007/978-3-031-91005-0_32)

- Pan, M., Wu, B., & Zhang, Y. (2020). Curriculum design of AI experimental course based on CDIO model. *Advances in Intelligent Systems and Computing*, 1075, 525–531. [https://doi.org/10.1007/978-3-030-32591-6\\_56](https://doi.org/10.1007/978-3-030-32591-6_56)
- Rutecka, P., Cicha, K., Rizun, M., & Strzelecki, A. (2025). Generative AI in curriculum design: Empirical insights into model performance and educational constraints. *IEEE Transactions on Learning Technologies*, 18, 757–768. <https://doi.org/10.1109/TLT.2025.3587081>
- Salhab, R. (2024). AI literacy across curriculum design: Investigating college instructors' perspectives. *Online Learning Journal*, 28(2). <https://doi.org/10.24059/olj.v28i2.4426>
- Selvakumar, P., Tamizharasu, S., Singh, K., & Lakshmi Prasanna, L. (2025). *Curricula and instructional design for AI* (pp. 17–35). <https://doi.org/10.4018/979-8-3693-8965-2.ch002>
- Su, J., & Zhong, Y. (2022). Artificial intelligence (AI) in early childhood education: Curriculum design and future directions. *Computers and Education: Artificial Intelligence*, 3, 100072. <https://doi.org/10.1016/j.caeai.2022.100072>
- Taylor, P. (2024). Blended learning challenges of EFL undergraduate students: Student learning experience in an AI-integrated ESP course. *Studies in English Language and Education*, 11(3), 1431–1449. <https://doi.org/10.24815/siele.v11i3.37472>
- Usman, N., Kusumanegara, A., Husnaini, H., Said, Y. R., & Ismayanti, D. (2025). *Cultural bias and linguistic representation in AI curriculum design: Implications for Society 5.0* (pp. 229–258). <https://doi.org/10.4018/979-8-3373-1062-6.ch008>
- Yin, D., Li, Q., Karimi, A., Mitsakos, N., El Mahbes, R., & Cox, J. (2024). *Integration of AI and domain expert knowledge to improve ESP efficiency and performance*. <https://doi.org/10.2118/221565-MS>