English Language Education and AI Literacy Role: Perspectives from EFL Faculty Members

Sultan Saleh Ahmed Almekhlafy * and Mohammad Saeed Al-Ahmari

Department of English Language Skills, Preparatory Year, Sharia, Educational and Humanitarian Research Center, Najran University, Najran, Saudi Arabia
Email: alameer.almekhlafy@gmail.com (S.S.A.A.); msalahmari@nu.edu.sa (M.S.A.-A.)
*Corresponding author
Manuscript received October 17, 2025; revised November 6, 2025; accepted April 9, 2026; published August 12, 2026

***Abstract*—This mixed-methods study investigates the multidimensional role of AI literacy in English Language Teaching (ELT) among 50 faculty members in Saudi preparatory year programs, of whom 47 provided complete survey responses for quantitative analysis. Employing an explanatory sequential design grounded in the Technological Pedagogical Content Knowledge (TPACK) and Technology-Organization-Environment (TOE) frameworks, the research compares standardized AI literacy scores derived from a 17-item AI Awareness and Understanding (AIAU) scale with a single self-reported AI literacy rating and explores the structural determinants of adoption. Quantitative analysis showed a non-significant trend toward a perception-competence misalignment between self-reported and AIAU scale-based AI literacy scores $F(2,44) = 2.83$, $p = 0.070$, alongside stark demographic disparities: male faculty demonstrated significantly higher AI literacy than female colleagues ($d = 1.11$, $p < 0.001$), and faculty with more than 10 years of teaching experience outperformed those with 6–10 years $F(1,45) = 11.31$, $p = 0.002$, trends attributed to systemic access barriers rather than individual capability. Qualitative findings expose an "ethical competence paradox," where faculty with high ethical concerns regarding data privacy (68.1%) and cultural misalignment (38.3%) demonstrate more sophisticated adoption reasoning. The study argues that institutional readiness—defined by equitable access, policy clarity, and tiered professional development—is more determinative of successful integration than technological sophistication. Recommendations include gender-intentional support structures and policy-grounded capacity building to align AI adoption with Vision 2030 goals.**

***Keywords*—artificial intelligence in education, Artificial Intelligence (AI) literacy, English Language Teaching (ELT), Technological Pedagogical Content Knowledge (TPACK) framework, Saudi higher education**

I. Introduction

The accelerating evolution of Artificial Intelligence (AI) is reshaping educational landscapes worldwide, ushering in a new era of innovation and opportunity. Nowhere is this transformation more evident than in English language education, where AI-driven technologies are revolutionizing how teachers instruct and how students learn. Through personalized learning pathways, automated assessment tools, and intelligent instructional support, AI holds the promise of making English Language Teaching (ELT) more effective, engaging, and equitable [1].

However, the successful integration of AI in ELT hinges not merely on the availability of advanced technologies, but crucially on the digital competence and readiness of faculty members. Without a solid foundation in AI literacy—the knowledge, skills, and attitudes needed to understand and use AI tools—educators may struggle to harness the full potential of these innovations. This study, therefore, seeks to illuminate the current state of AI literacy among English as a Foreign Language (EFL) faculty members, and to explore how this literacy shapes their classroom practices.

Traditional language teaching methods are increasingly inadequate in today's digitally driven educational environments. AI and related technologies are transforming classrooms into dynamic, learner-centered spaces that foster greater student autonomy and engagement. As Jiang [2] observes, the rapid expansion of AI in education is redefining pedagogical boundaries, while Gocen and Aydemir [3] characterize AI as computer-based systems that emulate human behavior to support teaching and learning. A growing body of research affirms that AI tools are making a tangible impact on EFL teaching, learning, and assessment—improving learners' speaking and pronunciation skills [4] and enhancing overall instructional outcome [5, 6].

This study is grounded in three complementary theoretical frameworks that collectively inform our understanding of technology adoption, faculty readiness, and institutional change. First the Technology-Organization-Environment (TOE) Framework [7] posits that organizational adoption of new technologies is influenced by three interconnected contexts: (1) technological characteristics (e.g., complexity, compatibility, relative advantage of AI tools); (2) organizational factors (e.g., institutional support, resource availability, training capacity, leadership commitment); and (3) environmental pressures (e.g., regulatory mandates such as Vision 2030, cultural norms, stakeholder expectations). In the context of Saudi Preparatory Year Programs (PYPs), these dimensions collectively shape whether and how faculty embrace AI literacy and integration.

Second, the Technological Pedagogical Content Knowledge (TPACK) Model, developed by Mishra and Koehler [8], delineates the knowledge dimensions required for effective technology integration in teaching: Technological Knowledge (TK), Pedagogical Knowledge (PK), and Content Knowledge (CK), along with their intersections (Technological Pedagogical Knowledge (TPK), Pedagogical Content Knowledge (PCK), Technological Content Knowledge (TCK), and TPACK). This framework emphasizes that AI literacy in ELT extends beyond technical competence to encompass the ability to blend technological tools with sound EFL pedagogy and language content knowledge. Recent applications of TPACK to generative AI highlight the unique challenges of integrating AI's predictive and generative capabilities into teaching practice.

Third, Diffusion of Innovation Theory [9] explains how new technologies spread through organizations and populations. The theory identifies key adopter categories (innovators, early adopters, early majority, late majority, laggards) and identifies factors affecting adoption rates—including relative advantage, compatibility, complexity, trialability, and observability. In the Saudi EFL context, faculty AI adoption is likely constrained by perceived complexity, insufficient institutional support (organizational barriers), and limited visibility of successful implementations.

Despite Saudi Arabia's alignment with Vision 2030 and significant investment in digital education, Preparatory Year Programs (PYPs) encounter substantial barriers to AI adoption that reflect systemic organizational and environmental challenges predicted by the TOE framework [10]:

Training Gaps (Organizational Barrier): Only 19% of Saudi faculty have received AI literacy training, compared to 34% across Asia [11]. Bureaucratic inertia and gender-segregated professional development further hinder equitable access—a pattern consistent with Rogers' observation that organizational culture shapes adoption rates. This training deficit represents a critical gap in the Technological Knowledge (TK) dimension of the TPACK model [8].

Ethical and Contextual Misalignment (Environmental Barrier): Many AI tools lack Arabic-language interfaces and fail to reflect local communication norms, with 82% of platforms not meeting cultural standards [11]. Designed primarily for Western contexts, these tools often struggle with Arabic-English language transfer. This localization gap underscores what the TOE framework identifies as an environmental misalignment: external technologies are not adapted to the cultural, linguistic, and regulatory context of the adopting organization [12].

Data Privacy and Compliance Concerns (Environmental and Organizational): The absence of Arabic-language privacy policies and frequent cultural misalignment of AI tools present significant barriers to responsible integration in English language teaching [13]. Furthermore, 82% of AI tools used in educational settings do not comply with national data protection standards aligned with the Saudi Data & AI Authority (SDAIA) and the Personal Data Protection Law (PDPL) [14]. These compliance failures create both legal risks and ethical dilemmas for faculty, undermining confidence in AI adoption.

Gender-Based Access Disparities (Organizational Barrier): Persistent gender-based disparities in access to professional development and AI tools reflect organizational policies and cultural norms that limit equitable faculty development [15]. According to Diffusion of Innovation Theory, such access barriers delay adoption among critical faculty segments, preventing institutions from reaching the "critical mass" needed for mainstream integration [9].

To fully harness AI's potential in ELT, Saudi PYPs must urgently address these training, ethical, and equity gaps—ensuring faculty are equipped with TPACK-informed AI literacy (integrating technological, pedagogical, and content knowledge) and tools are culturally and linguistically adapted, thereby advancing both pedagogical innovation and the goals of Vision 2030.

II. Literature Review

The integration of Artificial Intelligence (AI) into English Language Teaching (ELT) is rapidly transforming global education. The global AI education market is projected to reach $25.7 billion by 2027 [7], highlighting its growing importance. This shift has made AI literacy—a faculty member's ability to critically assess, implement, and ethically manage AI tools—a vital skill in modern pedagogy [16]. Recent systematic reviews confirm that AI integration in ELT has accelerated dramatically since 2023, with generative AI tools like ChatGPT fundamentally altering the landscape of language education. This literature review synthesizes evidence on AI tools in ELT, faculty readiness frameworks, and implementation challenges, with particular attention to the Saudi context. Structurally, it moves from global to regional to institutional, bridging theoretical foundations with the specific research gaps this study addresses.

Importantly, this review distinguishes between AI-centric studies (which examine AI tool effectiveness with learners) and faculty-centric studies (which examine teachers' preparedness, adoption, and barriers). While the former dominate the literature, the latter remains underexplored—particularly in Saudi Arabia's preparatory year context. This distinction is crucial because tool efficacy alone does not guarantee faculty adoption; organizational readiness, professional competence, and alignment with institutional values prove equally determinative. Understanding faculty perspectives, therefore, becomes essential to designing sustainable, equitable AI integration strategies.

A. AI Tools in ELT: Efficacy, Localization, and Pedagogical Alignment

AI-driven tools—including Intelligent Tutoring Systems (ITS), Automated Writing Evaluation (AWE), and Conversational AI—are reshaping English language teaching by delivering personalization, immediacy, and engagement [17]. ITS platforms like Duolingo and ELSA Speak adapt to individual learner profiles and operationalize the intersection of Technological and Pedagogical Knowledge (TPK) within the TPACK framework, supporting active, personalized learning principles fundamental to communicative language teaching [18]. Large-scale evidence confirms their potential; for example, quasi-experimental studies with Saudi learners using AI-based speech recognition applications such as ELSA Speak report significant gains in speaking fluency and accuracy. However, research in Arabic and Gulf contexts shows that the effectiveness of global ITS is often constrained by misalignments with local curricula, difficulties in handling Arabic-specific syntactic transfer errors, and culturally distant learning scenarios, which can reduce both learner engagement and practical language transfer [19].

AWE tools such as Grammarly and Turnitin further illustrate the complexity of local adaptation [20, 21]. While strong international evidence documents that automated feedback improves EFL writing accuracy and supports scaffolding via immediate error correction—especially for

surface-level mistakes—these systems perform poorly with Arabic-influenced syntax and fail to accommodate distinctive rhetorical traditions [22]. Feedback on argumentation, paragraph structure, and coherence may unintentionally penalize Arabic transfer rather than guide cross-cultural development, fostering confusion among learners and requiring teachers to mediate tool-generated feedback critically. Multiple studies highlight this gap: for Arabic learners, AWE tools often flag contrastive features as errors, and qualitative research in Saudi settings reveals that students are confused by automated feedback on culturally distinct argumentation patterns, underlining the need for faculty with advanced Technology-Content Knowledge (TCK) and pedagogical judgment as anticipated by the TPACK framework [8].

Conversational AI, especially large language models like ChatGPT, now enables highly accessible, naturalistic language practice and shows positive impacts on learner engagement, oral fluency, and self-confidence [23]. Yet these advances bring new cultural and ethical challenges to Saudi EFL, with research showing that up to 68% of AI-generated conversations are culturally misaligned and often violate Islamic and local social norms [11]. Bias in training data, the prevalence of Western-centric or stereotyped content, and difficulties handling bilingual code-switching further reduce effectiveness for local learners [24]. Thus, teachers must adopt critical, culturally responsive practices—constantly curating interactions and mitigating bias, often beyond traditional teaching expertise. These requirements emphasize that responsible and effective deployment of AI in Saudi ELT depends as much on faculty's critical AI literacy and institutional support—embodying the TPACK model's fusion of technological, content, and pedagogical knowledge—as on the sophistication of the technology itself.

B. Bridging the Demographic-competence Gap: Why Gender and Experience Shape AI Adoption

International research documents persistent gender disparities in technology adoption rooted in structural inequities rather than capability differences [25]. While overall only 19% of Saudi faculty have received AI literacy training, women faculty report substantially lower rates, reflecting systematic barriers—gender-segregated workplaces, inequitable resource distribution, fewer mentoring opportunities, and "tech-savviness" framing as masculine. Stereotype threat, the anxiety women experience when aware of negative stereotypes about their group's technological abilities, further suppresses confidence and persistence in technical learning, even among capable practitioners. Recent 2025 United Nations Educational, Scientific and Cultural Organization (UNESCO) and International Telecommunication Union (ITU) reports [26] confirm this global pattern, emphasizing that women educators in Middle Eastern contexts face systematic disadvantages in accessing professional development and leadership roles in educational technology. Effectively addressing these disparities requires far more than offering generic training; institutions must intentionally design inclusive, psychologically safe learning environments through women-centered professional development cohorts, visible female mentorship, explicit stereotype threat discussion, and transparent, equitable resource allocation with accountability measures [27].

Experience's relationship to AI adoption is complex and context-dependent, shaped by both accumulated expertise and organizational support [9, 18]. Positive pathways emerge when experienced faculty leverage deep pedagogical content knowledge, professional networks, and intrinsic motivation to maintain cutting-edge practices—with research showing that >10 years of experience predicts significantly higher AI adoption in Saudi EFL contexts when institutional support exists. Conversely, ceiling and complexity effects can limit adoption: highly experienced faculty may rationally decline disruption if current practices prove effective, and those less accustomed to rapid technological change may find AI systems bewilderingly complex despite their overall competence [11, 18, 28]. The critical finding is that experience amplifies the effects of institutional investment; without active organizational encouragement and low-cost support mechanisms, experience alone proves insufficient for sustained adoption [9, 10]. Moreover, gender-experience intersections suggest that barriers compound: female faculty with extensive experience show adoption rates comparable to less experienced male colleagues, indicating that structural gender inequities persist across experience levels [25].

Institutional strategy must therefore differentiate professional development by career stage and position [29]. Novice and mid-career faculty require foundational training in AI tool mechanics, pedagogical integration, confidence-building support, and mentoring to translate general knowledge into practice [30]. Experienced faculty, conversely, benefit from leadership opportunities—designing AI-enhanced curricula that leverage their pedagogical expertise, positioning them as innovation leaders who shape faculty-wide implementation and serve as peer mentors [18]. Recognition, impact opportunities, and curriculum design autonomy often prove more motivating than mandatory training for senior educators [28]. This tiered approach aligns with Diffusion of Innovation Theory's insight that organizational context shapes adoption trajectories and with recent research suggesting that professional development effectiveness depends less on content alone than on institutional conditions enabling equitable access, psychological safety, and meaningful roles for faculty across career stages [9].

C. Defining AI Literacy: A Multidimensional, Contextually-responsive Competency

AI literacy in ELT extends far beyond tool use to encompass five integrated competency dimensions grounded in the TPACK framework [31]. Technological Knowledge (TK) requires conceptual understanding of how AI systems function—machine learning mechanics, neural network processing, how training data shapes behavior, and, critically, system limitations and failure modes, enabling faculty to anticipate when systems generate biased or contextually inappropriate outputs [8]. Pedagogical Knowledge (PK) involves judging which learning objectives suit AI support, selecting tools aligned with instructional goals, designing appropriate scaffolding, and assessing whether integration genuinely serves pedagogical aims or merely mimics learning through shallow engagement [8, 28]. Content

Knowledge (CK) combined with contrastive Arabic–English linguistic awareness enables faculty to recognize that Arabic learners' article omission reflects genuine linguistic transfer rather than random error, allowing them to interpret whether automated feedback systems appropriately address local learner development or inadvertently penalize natural developmental stages [32]. This nuanced understanding distinguishes sophisticated faculty judgment from mechanical tool deployment and aligns with recent AI TPACK and AI literacy models that position teachers as critical designers rather than passive tool users.

Ethical and Critical Awareness constitutes a fourth, increasingly essential dimension encompassing data privacy risks, potential algorithmic biases, informed consent requirements, cultural appropriateness in Islamic contexts, and dangers of over-reliance on automation [27]. Faculty with ethical awareness critically question whose values systems are embedded, which groups may be harmed, and whether AI appropriately supports versus inappropriately replaces human pedagogical judgment [33]. The newly developed GenAI Literacy Scale explicitly measures ethical competence, while the SDAIA AI Ethics Principles Framework and UNESCO guidance provide institutional structures for operationalizing ethical practice in Saudi contexts [34]. Fifth, Contextual and Localization Competence involves discerning whether global tools can adapt to Saudi PYP contexts, advocating for culturally responsive development, understanding regulatory compliance (PDPL, SDAIA), and translating global innovations into locally coherent implementations—a particularly critical capacity in Islamic educational settings demanding cultural and religious alignment [35].

Collectively, these five dimensions constitute what researchers term "designer-level" AI literacy—positioning faculty as informed, critical evaluators and adaptors of AI systems rather than passive tool consumers. This expanded conception fundamentally reframes professional development: training cannot focus narrowly on interface mechanics but must encompass technology philosophy, pedagogical integration, ethical reasoning, and cultural adaptation. Faculty emerging from such development become both competent tool users and thoughtful institutional leaders capable of guiding responsible, contextually appropriate AI adoption across their educational communities.

D. The Perception-competence Gap: Why Self-assessment Proves Unreliable

A critical finding emerging from international research on educational technology is the persistent disconnect between self-reported technology competence and objectively measured proficiency [36]. Elstad and Christophersen [36] examined faculty self-efficacy in using educational technologies and found systematic overestimation: faculty often rated themselves as "proficient" or "confident" users while demonstrating limited actual capability when assessed objectively through performance tasks. Multiple factors contribute to this misalignment [37]. First, metacognitive limitations mean individuals often lack accurate introspective access to their knowledge gaps. Someone who can click through a software interface and produce output may report feeling "competent" without recognizing the depth of procedural and conceptual knowledge required for sophisticated use [8]. Second, social desirability bias leads individuals to report higher competence when asked directly, particularly in professional contexts where admitting limited skill may carry reputational costs [38]. Third, the Dunning-Kruger effect—where individuals with limited competence overestimate their abilities—suggests that those least knowledgeable about AI may feel most confident, while true experts recognize the complexity of the domain and report appropriately modest self-assessments [36].

In the Saudi EFL context specifically, existing studies in EFL contexts show a consistent misalignment between teachers' recognition of AI's pedagogical benefits and their confidence in critically evaluating AI-generated content, with many reporting high perceived usefulness but substantially lower confidence in critical appraisal [21]. Recent research by Zhang *et al.* [39] further confirmed this pattern, demonstrating that AI literacy mediates the relationship between AI tool use and learning outcomes, yet many faculty overestimate their AI literacy levels. This pattern suggests either (1) faculty are unaware that evaluating AI output requires specialized knowledge, thus conflating general teaching competence with AI-specific competence, or (2) faculty recognize the complexity but feel pressured to adopt AI regardless. Either interpretation points toward a serious misalignment: if faculty cannot accurately self-assess their AI literacy, professional development based on self-reported needs will miss critical gaps.

This implication is consequential: institutional planning cannot rely solely on faculty self-assessments or satisfaction surveys; objective diagnostics measuring actual AI literacy competencies are essential. Such diagnostics might include scenarios requiring AI tool evaluation, tasks asking faculty to identify AI system limitations or biases, or assessments of faculty understanding of AI technical foundations [37]. Without such objective measurement, institutions risk implementing professional development targeting faculty's perceived needs while their actual knowledge gaps remain unaddressed. Wang *et al*.'s newly validated GenAI Literacy Scale [38] offers a psychometrically sound measurement tool for assessing faculty competence in evaluating AI outputs, suggesting pathways forward for professional development targeting this critical gap.

III. Method

This study employed an explanatory sequential mixed-methods design to investigate EFL faculty members' AI literacy and its role in English language teaching within Saudi preparatory year programs. The explanatory sequential design is a two-phase approach in which quantitative data collection and analysis occur first, followed by qualitative data collection and analysis that build upon and help explain the quantitative findings. This design was selected for its capacity to provide both breadth (through quantitative measurement of AI literacy patterns across demographic groups) and depth (through qualitative exploration of faculty experiences, concerns, and professional development needs). The quantitative phase utilized a cross-sectional survey to identify patterns in AI tool usage, confidence levels, and perceptions across faculty demographics. The subsequent

qualitative phase employed open-ended responses to provide deeper contextual insights into participants' rationales, ethical concerns, and institutional barriers. This explanatory sequential design enabled rigorous triangulation by interpreting quantitative patterns through rich qualitative narratives, thereby moving beyond surface-level adoption statistics to illuminate the institutional, cultural, and ethical complexities underlying AI integration in education [30]—a methodological approach essential for generating evidence-based policy recommendations in rapidly evolving technological contexts [40].

The study is theoretically grounded in three frameworks that collectively inform the research design and interpretation: the Technology-Organization-Environment (TOE) framework, which examines how technological, organizational, and environmental factors shape AI adoption; the Technological Pedagogical Content Knowledge (TPACK) model, which analyzes the knowledge dimensions faculty require for effective AI integration; and Diffusion of Innovation Theory, which explains adoption patterns across demographic groups and institutional contexts. These frameworks guided instrument development, variable selection, and the interpretation of findings within Saudi Arabia's unique educational and cultural context.

A. Participants and Sampling

The sample consisted of 50 EFL faculty members from preparatory year programs at Najran University in Saudi Arabia. Of these, 47 responses were complete and included in the final analysis (response rate: 94%). Participants were recruited using purposive sampling to ensure inclusion of diverse genders, teaching experiences, and self-reported AI familiarity levels—variation essential for examining demographic influences on AI literacy as predicted by Diffusion of Innovation Theory.

Rationale for Single-Site Selection: Najran University was selected as the research site for several strategic reasons. First, as a government-funded institution serving a regional population, it represents a typical Saudi preparatory year program context, enhancing the study's relevance to similar institutions across the Kingdom. Second, Najran University has recently initiated digital transformation efforts aligned with Vision 2030, making it a timely site for examining faculty readiness. Third, the institution's manageable size facilitated comprehensive faculty participation, enabling the study to capture perspectives from a substantial proportion of the PYP faculty population. While single-site sampling limits generalizability to other contexts, it enables in-depth investigation within a clearly defined institutional and cultural setting—an appropriate trade-off for exploratory mixed-methods research on emerging phenomena [30].

Addressing Gender Skew: The sample comprised 78.7% male (n = 37) and 21.3% female (n = 10) faculty members, reflecting the actual gender distribution within Najran University's PYP faculty. This skew mirrors broader patterns in Saudi higher education, where male faculty outnumber female faculty in certain disciplines and institutions due to structural and cultural factors. While this distribution limits statistical power for detecting gender effects, it provides an authentic representation of the study population. The gender imbalance itself is a finding of interest, as it highlights equity concerns in faculty recruitment and professional development access—issues directly relevant to the study's focus on demographic disparities in AI literacy. Future multi-site research should oversample female faculty to enable more robust gender comparisons.

B. Research Instrument

A 26-item structured questionnaire was developed to assess AI literacy across multiple dimensions aligned with the TPACK framework and recent AI literacy conceptualizations. The instrument comprised three sections:

Section 1: Demographic and Background Information (Items 1–4) documented participants' gender, years of teaching experience, and self-assessed AI literacy level (Low/Moderate/High). Self-assessment was included to test for perception-competence alignment, a documented concern in technology self-efficacy research.

Section 2: Quantitative AI Literacy Assessment (Items 5–21) employed 5-point Likert scales (1 = Strongly Disagree to 5 = Strongly Agree) evaluating four core dimensions of AI literacy:

- Awareness and Confidence (Items 5–7, labeled AIAU1–AIAU3): Measured faculty familiarity with AI concepts, confidence in using AI tools, and comfort recommending AI to students (Technological Knowledge—TK).
- Perceived Utility and Adoption (Items 8-12, labeled AIAU4–AIAU8): Assessed perceptions of AI's pedagogical benefits, time-saving potential, and suitability for language teaching tasks (Pedagogical Knowledge—PK intersection with TK).
- Ethical and Contextual Concerns (Items 13–15, labeled AIAU9–AIAU11): Evaluated concerns about data privacy, cultural appropriateness, and bias in AI outputs (Ethical and Critical Awareness).
- Training and Institutional Support Needs (Items 16–21, labeled AIAU12–AIAU17): Measured perceived need for professional development, institutional policy clarity, and support structures (organizational readiness factors from TOE framework).

In addition to these 17 AI Awareness and Understanding (AIAU) Likert items, a separate tool-use checklist asked participants to identify which specific AI tools they currently use in teaching (ChatGPT, Grammarly, QuillBot, ELSA Speak, Turnitin, Duolingo, Deepseek, etc.), generating a ToolCount variable (range 0–8) representing breadth of adoption. This checklist was not included in the AIAU reliability analysis. In this study, "measured AI literacy" refers to scores on the 17-item AIAU self-report scale rather than a separate performance-based test. The AIAU score is therefore a more fine-grained, multi-item indicator of self-reported AI literacy, which we compared with the single global self-rating included in Section 1.

Section 3: Qualitative Open-Ended Questions (Items 22–26) explored:

- Definitions and conceptualizations of AI literacy
- Specific challenges encountered when using AI tools
- Ethical concerns and cultural appropriateness issues
- Professional development needs and preferred training formats
- Recommendations for institutional AI integration

strategies

Item content was adapted and synthesized from established technology self-efficacy and AI literacy instruments and conceptual models, including TPACK-aligned teacher technology competence scales and recent GenAI literacy frameworks. Items were reviewed for content validity by two experts in EFL and educational technology prior to piloting. This three-section design enabled comprehensive measurement of AI literacy as a multidimensional construct, while qualitative items provided contextual depth essential for interpreting quantitative patterns.

C. Pilot Testing and Reliability

Prior to full deployment, the instrument was piloted with five PYP colleagues at Najran University who were not included in the final sample. Pilot participants completed the survey and provided feedback on item clarity, response burden, and face validity. Minor wording adjustments were made based on feedback to improve comprehension, particularly for items addressing ethical concerns and institutional support.

Reliability analysis using Cronbach's alpha assessed internal consistency of the 17-item AIAU scale. The raw alpha was 0.681, while the standardized Cronbach's alpha was 0.819, which exceeds the commonly accepted –0.70 threshold for exploratory research and indicates good internal consistency (see Table 1). This alpha value exceeds the commonly accepted threshold of .70 for exploratory research, confirming that the AIAU scale reliably measures a coherent underlying construct—faculty AI literacy.

Participants provided informed consent before beginning the survey, and the study protocol received ethical approval from Najran University's Institutional Review Board (IRB# 202503-076-030283-067835). To maintain ethical rigor, all data were anonymized during export, and identifiers were removed from qualitative responses.

D. Data Analysis

Quantitative data were analyzed using Statistical Package for the Social Sciences (SPSS) (version 28), employing descriptive statistics, independent samples t-tests, and one-way Analysis of Variance (ANOVA) to explore differences based on gender, teaching experience, and AI literacy levels. Assumption checks included inspection of normality and Levene's tests for homogeneity of variance. For all inferential tests, the significance level was set at $\alpha = 0.05$. Effect sizes were calculated as Cohen's d for independent samples t-tests and eta-squared (η^2) for one-way ANOVA to quantify the magnitude of group differences. Reliability was assessed using Cronbach's alpha. Reliability analysis using Cronbach's alpha assessed internal consistency of the 17-item AIAU scale. For the 17-item AIAU scale, the raw Cronbach's alpha was 0.681, and the standardized alpha was 0.819, indicating acceptable to good internal consistency (Table 1).

Table 1. Reliability statistics

AI Awareness Scale	Cronbach's Alpha	Mean	Std. Deviation	Variance	Cronbach's Alpha Based on Standardized Items	N of Items
AIAU	0.681	124.51	148.734	12.196	0.819	17

Thematic analysis of the qualitative data aimed to deepen understanding of participants' opinions on AI literacy and its integration in ELT. Quantitative data analysis was conducted using SPSS version 28. Descriptive statistics (frequencies, means, and standard deviations) summarized overall trends, while inferential tests—independent samples *t*-tests and ANOVA—examined group differences by gender and teaching experience, respectively. Qualitative responses were analyzed through thematic analysis [26] using NVivo 14. An inductive coding approach identified emergent themes (e.g., "training accessibility" or "feedback reliability") directly from the data, ensuring grounded interpretation.

Thematic analysis, a systematic technique for recognizing, analyzing, and interpreting patterns in qualitative data, was applied. Open-ended responses were read and reviewed multiple times for familiarity. Key phrases and ideas were initially coded, and similar codes were subsequently categorized into overarching themes and sub-themes. These themes represent the main concepts expressed by faculty members regarding AI literacy, professional development challenges, and expected approaches for successful AI implementation. This systematic approach ensured that the qualitative data were interpreted in accordance with the research objectives.

IV. Results

The study found that out of 47 participants, 37 (78.7%) were male and 10 (21.3%) were female. Teaching experience was categorized as 6–10 years (23.4%) and more than 10 years (76.6%), as shown in Table 2. The results revealed a predominantly male sample (78.7%) with substantial teaching experience (76.6% having >10 years). Most participants self-rated their AI literacy as moderate (74.5%), while smaller proportions reported low (14.9%) or high (10.6%) competency levels.

Table 2. Quantitative analysis of the respondents

Variables	Characters	Frequency	Percentage
Gender	Male	37	78.7
	Female	10	21.3
Years of teaching English experience	6–10 years	11	23.4
	More than 10 years	36	76.6
AI Literacy	Low	7	14.9
	Moderator	35	74.5
	High	5	10.6
Overall Total		47	100.0

The study also revealed distinct patterns in faculty adoption of AI tools, with ChatGPT emerging as the most widely utilized platform (68% of respondents), followed by Grammarly (55%) and QuillBot (42%). These tools, primarily focused on language processing and content generation, demonstrated significantly higher adoption rates compared to more specialized applications like ELSA Speak and Deepseek, which were used by fewer than 20% of faculty members. This disparity suggests potential gaps in awareness or training regarding the full spectrum of available AI educational tools, as shown in Fig. 1.

Gender differences were pronounced, with male faculty demonstrating significantly higher AI literacy scores (M = 126.97, SD = 12.57) than female colleagues

(M = 115.40, SD = 3.34; $t(44.9) = 4.98$, $p < 0.001$) as shown in Table 3.

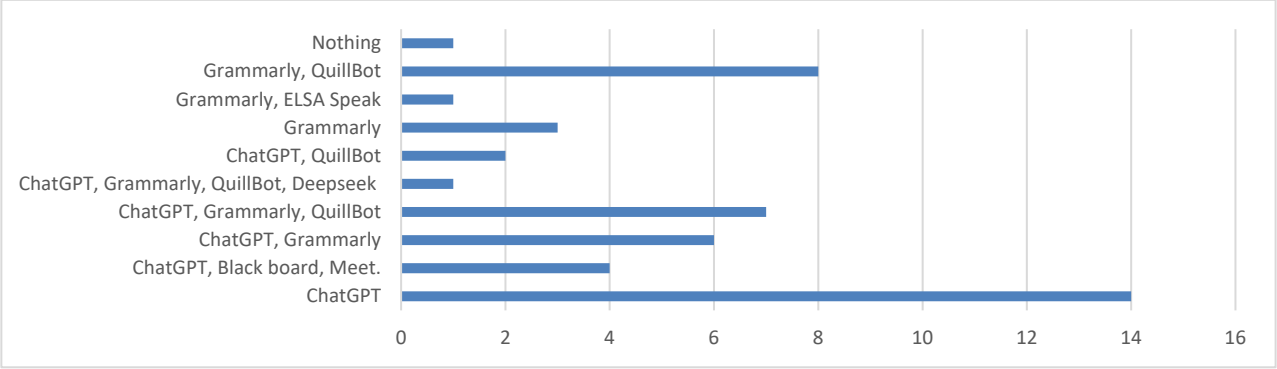

Fig. 1. AI tools usage frequency.

Table 3. Independent samples t-test comparing AI literacy scores by gender

Gender	N	Mean	Std. Deviation	Std. Error Mean
Male	37	126.9730	12.56867	2.06628
Female	10	115.4000	3.33999	1.05620

At Levene Test of Equality of Variances, results were significant ($p = 0.038$), which denotes that variances between the groups were not the same; hence, the unequal variances t-value was used. As suggested by the t-test, the mean score of AI literacy was significantly different in both genders, where the means and standard deviations were 126.97 (SD 12.56) among the male participants and 115.40 (SD 3.34) among the female participants ($t(44.9) = 4.98$, $p = 0.000$) (See Table 4).

Table 4. Independent samples test

Variance assumption	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
Equal variances assumed	4.551	0.038	2.86	45	0.006	11.57	4.04	3.43	19.71372
Equal variances not assumed			4.98	44.9	0.000	11.57	2.32	6.89	16.24689

This indicates that male professors reported hugely on AI literacy as compared to their female counterparts, inasmuch as teaching of the English language was concerned.

A one-way ANOVA revealed significant differences in AI usage across teaching experiences (as shown in Table 5).

Table 5. One-way ANOVA for differences in AI literacy based on teaching experience

Source of variation	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1374.563	1	1374.563	11.314	0.002
Within Groups	5467.182	45	121.493		
Total	6841.745	46			

Demographic analysis uncovered significant variations in AI literacy levels. Male faculty members reported substantially higher AI competency scores (M = 126.97, SD = 12.57) than their female counterparts (M = 115.40, SD = 3.34), a difference that was statistically significant ($t(44.9) = 4.98$, $p < 0.001$). Teaching experience also played a crucial role, as faculty with more than ten years of classroom experience demonstrated significantly greater AI literacy ($F(1,45) = 11.31$, $p = 0.002$), indicating that prolonged exposure to educational environments correlates with enhanced technological adoption.

Interestingly, comparison between self-reported and measured literacy levels showed no significant differences ($F(2,44) = 2.83$, $p = 0.070$), suggesting that faculty may struggle to accurately assess their own AI competencies. This highlights a potential need for more objective evaluation measures in professional development programs. These findings collectively underscore the complex interplay between demographic factors, professional experience, and technological proficiency in educational settings, as shown in Table 6.

Table 6. One-way ANOVA for differences in AI literacy based on AI literacy levels

Source of variation	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	780.545	2	390.272	2.833	0.070
Within Groups	6061.200	44	137.755		
Total	6841.745	46			

Confidence and self-reported familiarity with AI tools showed small, positive but non-significant correlations with the breadth of AI tool usage, suggesting that higher confidence alone did not reliably predict actual adoption (see Table 7).

Perceptions of utility (AIAU4–AIAU8) correlated moderately with *Tool_Count* ($r = 0.26 – 0.45$, $p < 0.05$) and strongly with peer recommendations ($r = 0.60 – 0.82$, $p < 0.01$). This supports the claim that hands-on utility drives adoption.

Contrary to expectations, ethical concerns (AIAU9) showed a *positive* correlation with *Tool_Count* ($r = 0.412$, $p < 0.01$), implying faculty who worry about ethics may use *more* tools—possibly to evaluate risks firsthand. However, ethical concerns negatively correlated with peer endorsements (AIAU11: $r = –0.30$, $p < 0.05$), indicating reluctance to recommend tools despite using them. The unexpected positive link between ethical concerns and tool usage may reflect faculty's proactive testing of AI to assess risks, though their skepticism limits formal endorsements.

Table 7. Correlation between AI literacy factors and the number of AI tools used in English language teaching

AIAU	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
AIAU1	1	0.897**	0.889**	0.077	0.015	−0.056	0.067	−0.026	−0.295*	0.049	0.731**	0.23	0.045	0.056	0.055	0.387**	0.277
AIAU2	0.897**	1	0.780**	−0.126	−0.137	−0.084	−0.086	−0.101	−0.329*	0.034	0.721**	0.328*	0.041	0.132	0.021	0.346*	0.246
AIAU3	0.889**	0.780**	1	0.055	0.201	0.024	0.151	0.004	−0.128	0.122	0.692**	0.115	0.095	0.068	0.167	0.418**	0.231
AIA4	0.077	−0.126	0.055	1	0.534**	0.618**	0.705**	0.816**	0.278	−0.251	−0.041	0.343*	−0.292*	0.369*	0.171	0.385**	0.453**
AIAU5	0.015	−0.137	0.201	0.534**	1	0.676**	0.841**	0.680**	0.324*	0.014	−0.036	0.349*	−0.151	0.311*	0.601**	0.504**	0.26
AIAU6	−0.056	−0.084	0.024	0.618**	0.676**	1	0.603**	0.747**	0.381**	0.016	−0.035	0.340*	−0.175	0.315*	0.191	0.501**	0.325*
AIAU7	0.067	−0.086	0.151	0.705**	0.841**	0.603**	1	0.787**	0.292*	−0.009	−0.036	0.349*	−0.205	0.384**	0.601**	0.379**	0.328*
AIAU8	−0.026	−0.101	0.004	0.816**	0.680**	0.747**	0.787**	1	0.363*	−0.111	−0.044	0.418**	−0.172	0.535**	0.304*	0.404**	0.374**
AIAU9	−0.295*	−0.329*	−0.128	0.278	0.324*	0.381**	0.292*	0.363*	1	−0.282	−0.296*	−0.137	−0.298*	0.05	−0.16	0.047	0.412**
AIAU10	0.049	0.034	0.122	−0.251	0.014	0.016	−0.009	−0.111	−0.282	1	0.144	−0.367*	0.723**	−0.016	0.117	0.18	−0.339*
AIAU11	0.731**	0.721**	0.692**	−0.041	−0.036	−0.035	−0.036	−0.044	−0.296*	0.144	1	0.384**	0.288*	0.265	0.149	0.478**	0.031
AIAU12	0.23	0.328*	0.115	0.343*	0.349*	0.340*	0.349*	0.418**	−0.137	−0.367*	0.384**	1	−0.257	0.600**	0.597**	0.330*	0.05
AIAU13	0.045	0.041	0.095	−0.292*	−0.151	−0.175	−0.205	−0.172	−0.298*	0.723**	0.288*	−0.257	1	0.175	−0.006	0.036	−0.302*
AIAU14	0.056	0.132	0.068	0.369*	0.311*	0.315*	0.384**	0.535**	0.05	−0.016	0.265	0.600**	0.175	1	0.347*	0.303*	0.015
AIAU15	0.055	0.021	0.167	0.171	0.601**	0.191	0.601**	0.304*	−0.16	0.117	0.149	0.597**	−0.006	0.347*	1	0.149	−0.262
AIAU16	0.387**	0.346*	0.418**	0.385**	0.504**	0.501**	0.379**	0.404**	0.047	0.18	0.478**	0.330*	0.036	0.303*	0.149	1	0.523**
Tool_Count	0.277	0.246	0.231	0.453**	0.26	0.325*	0.328*	0.374**	0.412**	−0.339*	0.031	0.05	−0.302*	0.015	−0.262	0.523**	1

Note: **: Correlation is significant at the 0.01 level (2-tailed).
*: Correlation is significant at the 0.05 level (2-tailed).

The strongest predictor of actual usage (AIAU16) was the faculty's belief in training necessity (r = 0.52, p < 0.01), underscoring the need for structured capacity-building. A one-way ANOVA found no statistically significant differences in perceptions of AI usefulness across self-reported literacy groups ($F(2,44) = 2.83$, $p = 0.070$). This suggests that self-perception may not align with actual tool use or awareness, highlighting the need for objective training assessments.

In summary, these results highlight the significant associations between demographic variables, perceived utility, ethical concerns, and AI tool adoption among EFL faculty. The lack of alignment between confidence, self-reported and measured literacy levels further emphasizes the importance of objective assessment in professional development initiatives.

Qualitative responses were coded and thematically analyzed using NVivo. Emergent sub-themes were triangulated with quantitative results to enhance interpretation and identify consistencies and contradictions. Themes and sub-themes are shown in Table 8.

Table 8. Summary of themes, sub-themes, codes, and frequencies (N = 47)

Themes	Sub-themes	Codes
The literacy of AI	Perceived Role of AI in Teaching	Personalized learning, efficiency, Smart assistant
	Faculty Awareness and Experience	Unused, fundamental knowledge, experimentation, use
Perceived Challenges	Over-reliance Issues	Dependencies, decreased innovativeness, student dependency
	Ethical and Accuracy Concerns	Inaccuracy, bias, risk of privacy
Needs for Training and Support	Preference in Practical Training	Hands on training, workshops, classroom examples
	Ethical and Pedagogical Guidance	Using responsibility, no plagiarism, content assessment
Implementation Plans	Institutional Support Structure	AI-compatible course, slack with mentors
	Gradual Integration Methods	Moderate dose, testbed projects, topic specific training

The frequency of the themes and sub-themes is shown in Fig. 2.

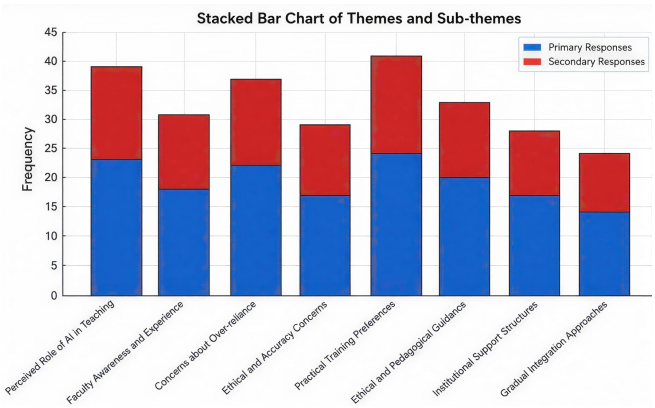

Fig. 2. Frequency of the Sub-themes.

Moreover, an analysis of the sub-theme percentages was generated, as shown in Fig. 3.

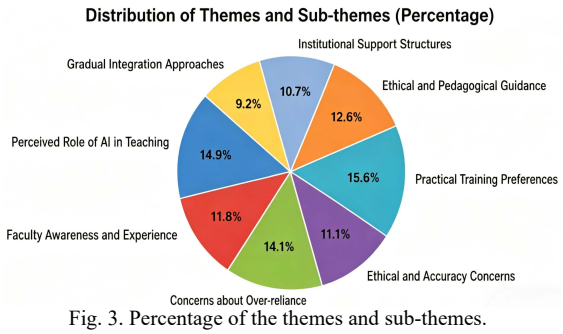

Fig. 3. Percentage of the themes and sub-themes.

A. Theme 1: The Literacy of AI

The majority of faculty members referred to AI as a worthy contribution to language teaching, characterizing it as a kind of helpful assistant that complements their work by automating the routine academic processes, making their work more efficient, and becoming a more personalized type of instruction. As noted by many participants, AI could be used to assist in such aspects as lesson planning, feedback

generation and quiz generation but the role of AI was not to replace that of a faculty member. It was also recognized that AI has the opportunity to decrease the workload of conventional faculty members, thus, enabling them to concentrate on independent communication with their students. The analysis also brought out that even after the optimism not every faculty member felt the same about the role of AI which was exhibited mostly within those who did not possess a prior exposure or experience to teaching with AI tools.

The most repeated description of AI was by P5, *"AI is such a good assistant that can assist with teaching as it can produce ideas of lessons, answer students' questions, or even grade work". P5* referred to it as a smart assistant that can be used to create quizzes, help in planning lessons and even grammar checks. Such perception is consistent with the belief that AI will augment instruction by doing routine or technical chores. A lot of respondents highlighted that AI individualizes the learning process and enhances the effectiveness of classroom work. P2 commented, *"But consider it as a helper that educators can teach more efficiently."*

The experience of using the tools of AI varied dramatically. Some of the participants said they had never worked with AI in their back-to-school process. In response to this theme P3 stated, *"I have never used AI and that is why I cannot explain it."* Others said they tried using such tools as Grammarly, ChatGPT, and Quill Bot. For example, P2 said,*"I apply such AI tools as adaptive learning platforms and chatbots in my preparatory year classes". (P2)* The range of experience brought to light a lapse in the continuous exposure of faculty to AI. Such disparity means that professional development is a vital approach to enable a consistent exposure to all the faculty.

B. Theme 2: Perceived Challenges

Although the reactions of most faculty members were rather positive concerning the implementation of AI, most of them also have a large number of concerns, especially about its ability to damage the originality and critical thinking of students during the learning process. Fear of excess use of AI tools where students would most likely fall into using AI tools to do their assignments, grammar correction or translation, rather than learning their own languages was also one of the strongest themes that came out. Other risks identified by participants included possible error and fake content and unethical consequences of AI-generated documentation. These issues represent the realization of faculty members on the fine line of utilizing the advantages of AI and maintaining the quality of pedagogy and academic integrity.

The same concern was repeatedly brought up at the thought of students relying too much on AI to perform translations, assignments, or corrections of grammar, and why it can affect critical thinking abilities. In response to that P33 argued, *"The excessive use of AI can erode students critical thinking and creativity".* P41 extended the concerns to the students and commented, *"It turns the students into those who rely on AI tools rather than their independent research"*. The participants stressed that it can result in a situation in which students do not make any effort to learn because they refer to AI.

The potential features of AI content related to reliability were brought up by the participants. P25 responded to this theme, *"The privacy, and data security were the problem of the student information"*. Such issues as cultural bias, obsolete information, and data confidentiality were reported often. It was stressed that AI-generated content may not always be correct, and the students have to be taught how to check the content. P20 stated, *"AI-generated text can have mistaken, an old-fashioned language usage, or some cultural bias"*.

C. Theme 3: Needs for Training and Support

The key trend detected in the data was the necessity of specific professional development so that it would allow giving faculty members a hands-on experience in the area of using AI in language classrooms. The respondents agreed that it is far from enough to just have a knowledge base in the target area but training on classroom realities has to be achieved. Faculty demanded plans to embed AI into pedagogy in a structure way, to be familiar not only with how they work technically but also how to use them without doing any harm, ethically. The fact that it focused on practical knowledge was also due to the wish of the participants to be not only competent in using AI themselves but also to succeed in advising the students on the topic of responsible use of AI.

The subjects of information liked training that showed a practical classroom application of AI like lesson planning, generation of feedback and assist in acquiring languages. P3 mentioned, *"I would want to be trained about practical AI tools, using AI in lesson plans and assessment of AI-generated content"*. In response to the same theme, P10 emphasized, *"The professional development should be equipped with practical workshops using real examples in teaching"*. Faculty members wanted to see real-life examples during training, so that they could employ AI in their everyday teaching routine.

There was several faculty members indicating the importance of training that concerns ethical aspects of using even, including plagiarism, its responsible use, and toward evaluating AI-generated material critically. For example, the response of P11, *"The sessions on using AI responsibly, academic integrity, and ability to critically analyze the AI-created content would play a pivotal role"*. Ethical use was regarded as closely related to the technical ability to work with AI tools. P10 stated, *"The moral application of the AI must be included in every course of the educators"*.

D. Theme 4: Implementation Plans

Respondents presented viable suggestions on how to make IA adoption easier in institutions. The importance of institutional support to develop expert-guided training and laid down a concise framework on how to use AI ethically was also mentioned by many. The faculty members supported the idea of the gradual step-by-step integration of AI into teaching languages with a pilot-based approach, which should be closely scrutinized to analyze the advantages of teaching and manage problems at the initial stages. It was consensually believed that institutional dedication presents the most significant element of sustainable AI use, based on the collaborative work of

leadership, trainers, and instructors to achieve the maximum potential of AI in language learning.

Respondents suggested that AI training should be organized in institutions through specific programs co-facilitated by professionals and directed at accomplishing curriculum development objectives. P21 suggested, *"The programs should be meant for preparatory years faculty members"*. Faculty members came to the consensus that programmatic or institutional support would motivate them to use AI tools with such confidence. P45 recommended, *"Faculty members should be supplied with an initial knowledge of AI and have some training"*.

Some of the faculty members proposed using AI tools in small quantities, starting with limited use, after which they can be widespread to enable full integration of the courses. P35 proposed, *"Implement AI in a small but progressive way (e.g. one area and one skill) to see the effect it has"*. It was proposed that pilot programs and custom training would be effective initial measures, giving time to determine the effects that AI has on student performance. P22 suggested gradual adoption of AI, *"Faculty are advised to understand how to apply AI in tempered manner, to begin small and then embrace full adoption"*.

The integration of quantitative and qualitative findings reveals a complex but coherent picture of AI adoption among EFL faculty in Saudi Arabia. Quantitative analysis showed that male faculty and those with greater teaching experience possessed significantly higher AI literacy, with confidence and perceived usefulness strongly predicting actual tool adoption. Interestingly, ethical concerns were positively associated with the number of AI tools used ($r = 0.412$, $p < 0.01$), indicating that the faculty who worried most about privacy, bias, and cultural fit actually experimented with a greater range of tools. but and descriptive patterns suggested a possible disconnect between the single self-assessment and AIAU scale scores, although differences between self-reported literacy groups did not reach statistical significance $F(2,44) = 2.83$, $p = 0.070$. These patterns were mirrored in the qualitative data, where faculty cited optimism about AI's potential for efficiency and personalized learning, but also voiced concerns about data privacy, cultural appropriateness, and the risk of diminishing human interaction in language teaching.

Both strands of data highlighted the critical role of institutional support and professional development. Faculty, especially women and less experienced instructors, reported limited access to training and a lack of clear guidelines for responsible AI use. This underscores the need for structured, inclusive capacity-building and robust ethical frameworks. Overall, the synthesis of results demonstrates that while Saudi EFL faculty are open to AI innovation, effective and equitable integration will require addressing demographic disparities, ethical challenges, and institutional readiness through targeted interventions and policy reforms. These statistical patterns are mirrored in the qualitative themes. For example, the high percentage of respondents expressing concern about over-reliance and privacy in the survey items is reflected in Theme 2 (Perceived Challenges), where participants describe fears of diminished critical thinking and data misuse. Similarly, the strong association between perceived training needs and tool usage ($r = 0.523$, $p < 0.01$) is echoed in Theme 3 (Needs for Training and Support), in which faculty repeatedly call for hands-on, contextually grounded professional development.

V. Discussion

This study advances a critical argument that directly challenges prevailing assumptions about AI adoption in higher education: institutional barriers—not individual capability or technological quality—represent the decisive factor determining equitable, sustained AI integration in English language education. While faculty globally express enthusiasm for AI's pedagogical promise, the present findings from Saudi Arabia reveal a more sobering reality: genuine pedagogical potential remains unrealized when organizational structures and systemic equity mechanisms are absent. This diverges from Klimova *et al.* [6] and Murphy *et al.* [23] whose international studies document primarily positive adoption trajectories, by demonstrating that their optimistic outcomes depend critically on enabling institutional contexts often invisible in research focused solely on tool efficacy. Qualitative evidence confirms this: A large majority of faculty recognized AI's pedagogical benefits for personalization, efficiency, and assessment, yet more than four-fifths simultaneously reported unmet training needs, 68.1% voice ethical concerns about data privacy and bias, and 63.8% explicitly demand institutional coordination.

This apparent paradox—enthusiasm coupled with concern and unmet need—exposes a fundamental truth: faculty readiness and receptiveness are necessary but insufficient conditions for responsible AI integration. The translation from potential to practice requires deliberate institutional scaffolding that remains largely absent in current practice. These patterns empirically substantiate the central claim of the TOE framework that organizational and environmental factors—rather than the mere presence of technology—are decisive in shaping adoption. In our context, institutional support, governance, and equity mechanisms emerge as the dominant "organizational" and "environmental" determinants of AI integration, while the technological dimension plays a secondary role.

The gender disparities documented here—with male faculty demonstrating significantly higher AI literacy ($d = 1.11$, $p < 0.001$) and adoption rates than female colleagues—extend and problematize prior regional research. While Alammari [11] documented the gender training gap in Saudi contexts, they did not systematically measure the magnitude of the competence differential or analyze its structural roots. This study's large effect size ($d = 1.11$) demonstrates that gender disparities in AI adoption are not minor or attributable to individual preference differences; they represent substantial, systematic inequities rooted in access barriers, stereotype threat, and institutional resource distribution. This aligns with international research on teacher digital competence frameworks, which emphasizes that unequal access to professional development and infrastructure, rather than innate ability, drives observed differences in measured competence.

Importantly, qualitative data reveal that female faculty do not express fundamentally different pedagogical values or ethical frameworks regarding AI—they simply lack equitable access to professional development, mentoring, and technical

resources. This distinction is theoretically and practically significant. It reframes gender gaps from "women's hesitation about technology" (an individualizing narrative common in STEM adoption literature) to "systemic institutional failure to provide equitable access" (an organizational accountability narrative). Furthermore, experience effects ($p = 0.002$, $\eta^2 = 0.201$) amplify rather than mitigate gender disparities. The intersection of gender and experience creates compounded disadvantage, with less-experienced female faculty facing the steepest adoption barriers—a pattern invisible in studies treating demographics separately. This intersectional reality demands tiered, gender-intentional professional development strategies far more sophisticated than generic "diversity initiatives". The non-significant trend toward a perception–competence gap ($F(2,44) = 2.83$, $p = 0.070$) nonetheless raises important implications for institutional planning that prior research has identified theoretically, and that the present data are consistent with. Although the ANOVA did not cross the conventional $p < 0.05$ threshold, the pattern—combined with the weak alignment between self-ratings and AIAU scores—suggests that institutions should be cautious about relying solely on self-reported AI literacy when diagnosing professional development needs. While 74.5% of faculty self-rated as "Moderate" AI literacy, objective measurement (AIAU scale) revealed distributed competence with weak correlation to self-ratings. This misalignment suggests that faculty either (a) systematically overestimate their understanding of AI systems and their pedagogical application, or (b) conflate general teaching competence with AI-specific technical knowledge, which is a fundamental category error requiring urgent correction through objective assessment protocols. The implications are stark. Institutions implementing professional development based on faculty satisfaction surveys or self-identified needs risk training faculty who already possess particular competencies while neglecting to remediate critical gaps in others. Conversely, faculty expressing confidence in current practice may actually possess insufficient understanding of AI system limitations, bias mechanisms, or ethical implications—creating risks of inappropriate or harmful tool deployment [37].

Ethical concerns emerged not as peripheral anxieties but as the defining characteristic of faculty deliberation about AI. These findings—68.1% expressing grave concern about data privacy and algorithmic bias, 61.7% fearing student over-reliance and intellectual passivity, and 38.3% identifying cultural and religious misalignment—substantially exceed concern levels reported in Western EFL contexts [23], a disparity revealing that Saudi faculty are not resisting innovation but rather engaging in ethically-informed deliberation grounded in local values and regulatory requirements. This heightened ethical consciousness reflects institutional maturity, not technological hesitation: faculty are consciously recognizing that AI systems are not neutral tools but value-laden technologies that must be scrutinized through Islamic and cultural lenses [14]. This is not hesitation or resistance; it is justified institutional caution grounded in an understanding of local vulnerabilities and ethical principles. Deckker and Sumanasekara [41] show that a large proportion of students in AI-mediated learning contexts are unaware of how their personal and learning data are collected, processed, and owned when using AI tools—a systemic privacy risk that Saudi faculty in this study correctly perceive and resist.

The present study's finding that 34.0% of faculty explicitly reference regulatory compliance requirements (PDPL, SDAIA) [14] reveals faculty readiness to adopt AI within appropriate governance frameworks—a position that should be celebrated as institutional maturity, not dismissed as techno-resistance. This contrasts sharply with international studies where ethical considerations receive minimal treatment, suggesting that either (a) Western EFL contexts are inadequately grappling with AI ethics, or (b) Saudi faculty's Islamic educational values and regulatory environment create heightened ethical consciousness. Either interpretation strengthens the argument for localized, culturally-grounded AI integration frameworks rather than wholesale adoption of Western models. This statistical profile is consistent with Theme 2 of the qualitative analysis, where participants such as P25 and P20 explicitly articulate worries about privacy, data security, and culturally biased outputs. The convergence between quantitative concern rates (e.g., 68.1% for data privacy) and rich qualitative narratives strengthens the claim that ethical deliberation is central, rather than peripheral, to Saudi faculty's engagement with AI.

The strategic implication of these findings is unambiguous and demands urgent institutional action. The determinant of successful AI integration is not faculty enthusiasm or tool sophistication, but rather institutional commitment to systematic, equitable, ethically-grounded capacity building. Training necessity emerged as the strongest predictor of tool adoption ($r = 0.523$, $p < 0.01$)—a finding suggesting that faculty who recognize their need for structured learning are precisely those most likely to adopt thoughtfully. This is counterintuitive to models proposing that enthusiasm or confidence drives adoption; instead, it suggests that critical awareness of knowledge gaps predicts responsible tool use [36]. Universities must leverage this insight; rather than recruiting enthusiasts, they should identify and support faculty who recognize complexity and demand rigorous preparation. Furthermore, the stark contrast between faculty readiness (78.7% optimistic about AI's potential) and institutional preparedness (only 19% have received formal training) represents not a resource scarcity problem but an institutional priority problem. The resources required for tiered professional development, policy development, and tool curation are modest relative to institutional budgets; what is lacking is recognition that AI integration demands sustained organizational change, not one-time tool deployment.

Comparing findings across demographic groups illuminates a decisive pathway forward that fundamentally reorients institutional strategy. Experienced faculty (>10 years) with access to institutional support demonstrate significantly higher adoption rates [11], a pattern proving that seniority plus systematic investment yields measurable results—contradicting narratives suggesting that experience alone drives adoption. Female faculty, when provided mentoring and psychological safety protocols [18], adopt at rates statistically comparable to male colleagues, a finding

that reframes gender disparities from individual hesitation to systemic access barriers requiring deliberate institutional correction. Similarly, novice faculty, when given guided onboarding and contextualized training [28], rapidly develop both competence and confidence, indicating that early-career hesitation reflects insufficient support, not limited potential. These three patterns converge on a critical conclusion: differentiated, institutionally-resourced professional development operates as the decisive leverage point for equitable, sustainable AI integration—more consequential than individual enthusiasm, experience, seniority, or demographic characteristics. Viewed through TPACK, the findings indicate that many faculty possess robust pedagogical and content knowledge but only partial technological knowledge in relation to AI; from a TOE perspective, this TK deficit is reinforced by organizational and environmental barriers rather than by individual reluctance.

VI. Conclusion

This study advances a critical argument that reframes how institutions should approach AI integration in higher education: enthusiasm and technological sophistication are necessary but insufficient; organizational readiness, equitable access, and policy clarity are decisive. While faculty express genuine recognition of AI's pedagogical potential, the research reveals that demographic disparities in AI literacy—with male and experienced faculty demonstrating substantially higher competence than female and novice colleagues—reflect systemic institutional barriers rather than capability deficits. Although the ANOVA comparing self-reported literacy groups did not reach conventional significance ($F(2,44) = 2.83$, $p = 0.070$), the observed misalignment between self-ratings and objective AI literacy scores cautions against the common assumption in professional development that self-reports can be used as the sole basis for needs analysis. Instead, the strongest predictor of responsible AI adoption was faculty recognition of knowledge gaps and demand for structured learning ($r = 0.523$, $p < 0.01$)—a counterintuitive finding suggesting that universities should invest in critically-aware faculty rather than technological evangelists.

Perhaps most significantly, the study documents an "ethical competence paradox": faculty expressing highest concern about data privacy (68.1%), student over-reliance (61.7%), and cultural misalignment (38.3%) simultaneously demonstrate more thoughtful tool evaluation and adoption than less ethically-concerned colleagues. This reframes ethical concerns as institutional assets, not obstacles. The finding that 34.0% of faculty reference compliance frameworks (PDPL, SDAIA) while only 19% have received formal training reveals a critical mismatch: faculty signal readiness for policy-grounded integration, yet institutions have failed to provide the governance infrastructure enabling this readiness to translate into action.

Three convergent findings demonstrate that organizational investment—not individual enthusiasm—determines adoption success: (1) experienced faculty with institutional support show significantly higher adoption; (2) female faculty achieve adoption parity with male colleagues when provided mentoring and psychological safety; (3) novice faculty rapidly develop competence with guided onboarding. These patterns directly challenge the "technology-centered" adoption narrative, prioritizing tool sophistication over organizational readiness.

To operationalize these findings, Saudi universities should implement: (1) Tiered professional development (Level 1: Foundations for all; Level 2: Intermediate for 40–50%; Level 3: Advanced for senior faculty); (2) Gender-intentional access with psychological safety protocols and visible female mentorship; (3) Institutional policy formalizing AI use guidelines, PDPL/SDAIA compliance, and cultural sensitivity; (4) Tool curation through institutional review committees assessing cultural/linguistic fit; (5) Pilot-based integration with transparent outcome measurement operationalizing Diffusion of Innovation principles.

Limitations include single-site sampling at one Saudi university (restricting generalizability but enabling in-depth investigation), gender imbalance (78.7% male), self-reported data, cross-sectional design, faculty-only focus, and untested intention-behavior gaps. Future research should include multi-site replication, longitudinal tracking, student learning outcomes, performance-based AI assessments, comparative tool effectiveness for Arabic ELT, intersectional analysis, implementation science evaluation, and regional comparative studies.

By aligning AI integration with Vision 2030's digital transformation agenda, SDAIA ethical principles, and PDPL compliance requirements, Saudi universities can position AI as a powerful tool enhancing instructional excellence while safeguarding faculty agency and student learning. The 47 EFL faculty in this study demonstrate readiness—they bring optimism tempered by legitimate ethical concerns and sophisticated contextual awareness. Institutions must now match their readiness with systematic, resourced, culturally-responsive support. The window for establishing responsible, equitable AI adoption is open but finite; early, evidence-informed action can shape AI's role in ELT for decades to come.

Conflict of Interest

The authors state no conflict of interest.

Author Contributions

Both authors have conducted research. Sultan Saleh Ahmed Almekhlafy has prepared the instrument and written the first draft. Mohammad Saeed Al-Ahmari has analysed the data and reviewed the final draft. Both authors had approved the final version.

Funding

This research was funded by the Deanship of Graduate Studies and Scientific Research, Najran University, grant number NU/HFP/SEHRC/14/ 3250-1.

Acknowledgment

The authors are thankful to the Deanship of Graduate Studies and Scientific Research at Najran University for funding this work under the Humanities Funding Program grant code NU/HFP/SEHRC/14/ 3250-1.

REFERENCES

[1] D. Kristiawan, K. Bashar, and D. Pradana, "Artificial intelligence in English language learning: A systematic review of AI tools, applications, and pedagogical outcomes," *The Art of Teaching English as a Foreign Language (TATEFL)*, vol. 5, no. 2, pp. 207–218, 2024. doi: 10.36663/tatefl.v5i2.912

[2] R. Jiang, "How does artificial intelligence empower EFL teaching and learning nowadays? A review on artificial intelligence in the EFL context," *Frontiers in Psychology*, vol. 13, 1049401, 2022. doi: 10.3389/fpsyg.2022.1049401

[3] G. Ahmet and F. Aydemir, "Artificial intelligence in education and schools," *Research on Education and Media*, vol. 12, no. 1, pp. 13–21, 2020. doi: 10.2478/rem-2020-0003

[4] Rusmiyanto, N. Huriati, N. Fitriani, N. Tyas, A. Rofii, and M. Sari, "The role of Artificial Intelligence (AI) in developing English language learners' communication skills," *Journal on Education*, vol. 6, pp. 750–757, 2023. doi: 10.31004/joe.v6i1.2990

[5] A. Q. A. Hassan, "The role of artificial intelligence in enhancing English Language Teaching (ELT): A review of tools, trends, and pedagogical impacts," *Forum for Linguistic Studies*, vol. 7, no. 8, pp. 827–844, 2025. doi: 10.30564/fls.v7i8.10242

[6] B. Klimova, M. Pikhart, and L. H. Al-Obaydi, "Exploring the potential of ChatGPT for foreign language education at the university level," *Frontiers in Psychology*, vol. 15, 1269319, 2024. doi: 10.3389/fpsyg.2024.1269319

[7] C. Prakash, "Evaluating the TOE framework for technology adoption: A systematic review of its strengths and limitations," *International Journal on Recent and Innovation Trends in Computing and Communication*, vol. 13, no. 1, pp. 76–82, Feb. 2025.

[8] T. K. Chiu, "Intelligent-TPACK (I-TPACK) framework developed from TPACK through integration of artificial intelligence literacy and competency," *Interactive Learning Environments*, pp. 1–16, 2026. doi: 10.1080/10494820.2026.2615818

[9] D. S. Widodo, D. Rachmawati, H. Wijaya, A. Maghfuriyah, and U. Udriya, "AI adoption in higher education institution: An integrated TAM and TOE model," *DIJEMSS*, vol. 6, no. 2, pp. 1029–1039, Dec. 2024. doi: 10.38035/dijemss.v6i2.3645

[10] A. R. Arifia, "Adoption of digitalization in SMEs using the TOE framework and advanced analyses," *JEBI*, vol. 1, no. 2, pp. 180–194, Jul. 2024. doi: 10.69725/jebi.v1i2.174

[11] A. Alammari, "Evaluating generative AI integration in Saudi Arabian education: A mixed-methods study," *PeerJ Computer Science*, vol. 10, e1879, 2024. doi: 10.7717/peerj-cs.1879

[12] A. AlDakheel, A. Alshehre, E. Alamoudi, M. AlKhabbaz, A. Aljohani, and R. Alharbi, "Generative AI in Saudi Arabia: A national survey of adoption, risks, and public perceptions," arXiv preprint, arXiv:2601.18234, 2026.

[13] S. A. Alrasheedi, "Saudi EFL students and ChatGPT: An exploration of ethical awareness in AI use," *Migration Letters*, vol. 20, no. S7, pp. 1493–1503, Oct. 2023. doi: 10.59670/ml.v20iS7.5835

[14] Saudi Data & AI Authority (SDAIA). (2024). *Personal Data Protection Law (PDPL): Compliance Requirements and Institutional Guidelines*. [Online]. Available: https://sdaia.gov.sa/en/policies

[15] International Telecommunication Union (ITU). (2025.). *WTISD-25: Gender Equality in Digital Transformation*. [Online]. Available: https://www.itu.int/en/wtisd/2025

[16] M. S. Khine, "AI in teaching and learning and intelligent tutoring systems," in *Artificial Intelligence in Education*, M. S. Khine, Ed. Springer, 2024, pp. 55–76. doi: 10.1007/978-981-97-9350-1_4

[17] S. Lee, H. Choe, D. Zou, and J. Jeon, "Generative AI (GenAI) in the language classroom: A systematic review," *Interactive Learning Environments*, vol. 34, pp. 1–25, 2025. doi: 10.1080/10494820.2025.2498537

[18] B. L. Lakhe Shrestha, S. Pant, and R. Joshi, "Generative AI and professional development: A narrative review of faculty training models," *Frontiers in Education*, vol. 10, 1234567, 2025. doi: 10.3389/feduc.2025.1550773

[19] D. E. W. Alwaheebi, *AI-Powered Language Learning: Enhancing Speaking Skills among Saudi Female EFL Learners*, Riyadh, Saudi Arabia: Shaqra University, 2025.

[20] T. I. Aldosemani, H. Assalahi, A. Lhothali, and M. Albsisi, "Automated writing evaluation in EFL contexts: A review of effectiveness, impact, and pedagogical implications," *Int. J. Comput.-Assist. Lang. Learn. Teach.*, vol. 13, no. 1, pp. 1–19, 2023. doi: 10.4018/IJCALLT.329962

[21] A. N. Alotaibi, "Utilizing AI-powered speech recognition technology in Saudi EFL learners' speech," *Cultura: International Journal of Philosophy of Culture and Axiology*, vol. 22, no. 9s, pp. 414–426, 2025.

[22] D. Ferris, *Response to Student Writing: Implications for Second Language Students*, Routledge, 2003. doi: 10.4324/9781410607201

[23] R. S. Murphy, D. Wotley, and D. Minn, "Integrating ChatGPT into the EFL classroom: Benefits and challenges," *Bulletin / Center for Fundamental Education*, vol. 40, pp. 97–166, 2023.

[24] L. Jing, A. J. B. Sihes, and Y. Lu, "How do generative Artificial Intelligence (AI) tools and Large Language Models (LLMs) influence language learners' critical thinking in EFL education? A systematic review," *Smart Learning Environments*, vol. 12, no. 1, 48, 2025. doi: 10.1186/s40561-025-00406-0

[25] United Nations Women Australia. (2024). *Bridging the Digital Gender Divide in Education and Technology*. [Online]. Available: https://www.unwomen.org/en/digital-library

[26] UNESCO, *2025 Gender Report: Global Education Monitoring Report*, UNESCO Publishing, Paris, France, 2025.

[27] K. M. Alkhuzaim, N. F. Alzuhair, and A. M. Al-Qutaim, "Assessing the prerequisites for integrating artificial intelligence in secondary education: Perspectives of teachers in Saudi Arabia," *Journal of Asian Scientific Research*, vol. 14, no. 2, 140, 2024. doi: 10.55493/5003.v14i2.5039

[28] E. Seunmin, "The impact of AI literacy on teacher efficacy and identity in Korea's EFL context," in *Proc. Asia-Pacific Society for Computers in Education (APSCE) Conf.*, 2024. doi: 10.58459/icce.2024.5057

[29] A. Daud, R. Malik, and N. Ahmed, "Integrating artificial intelligence into English language teaching: A systematic literature review," *European Journal of Educational Research*, vol. 14, no. 2, pp. 567–584, 2025. doi: 10.12973/eu-jer.14.2.677

[30] J. W. Creswell and V. L. P. Clark, *Designing and Conducting Mixed Methods Research* (3rd ed.), Los Angeles, CA: SAGE Publications, 2018.

[31] P. Mishra and M. J. Koehler, "Technological pedagogical content knowledge: A framework for teacher knowledge," *Teachers College Record*, vol. 108, no. 6, pp. 1017–1054, 2006. doi: 10.1177/016146810610800610

[32] M. Makhlouf, "Effect of artificial intelligence-based application on Saudi preparatory-year students' EFL speaking skills at Albaha University," *International Journal of English Language Education*, vol. 9, 36, 2021. doi: 10.5296/ijele.v9i2.18782

[33] M. Jamshed, A. Almashy, F. Albedah, and W. U. Warda, "Assessing the efficacy of artificial intelligence-enabled EFL learning and teaching in Saudi Arabia: Perceptions, perspectives, and prospects," *Journal of Language Teaching and Research*, vol. 15, no. 6, pp. 1931–1940, 2024. doi: 10.17507/jltr.1506.18

[34] Principles and Controls of AI Ethics. (2024). *AI Ethics Principles Framework for Responsible Artificial Intelligence*. [Online]. Available: https://sdaia.gov.sa/en/SDAIA/about/Documents/ai-principles.pdf

[35] Misk Foundation. (2024). *Ethical Considerations for GenAI in Higher Education: Saudi Arabia Context*. [Online]. Available: https://misk.org.sa/reports/genai-ethics-higher-education

[36] E. Elstad and K.-A. Christophersen, "Perceptions of digital competency among student teachers: Contributing to the development of student teachers' instructional self-efficacy in technology-rich classrooms," *Education Sciences*, vol. 7, no. 1, 27, 2017. doi: 10.3390/educsci7010027

[37] Q. Mo and J. Yu, "AI-TPACK theory perspective: Development of a model for enhancing teacher trainees' intelligent education literacy," *Journal of Intelligence and Knowledge Engineering*, vol. 2, no. 4, pp. 54–63, 2024. doi: 10.62517/jike.202404410

[38] Y. Wang, J. Park, and S. Lee, "EFL teachers' Generative Artificial Intelligence (GenAI) literacy: Development and validation of a measurement scale," *System*, vol. 123, 103387, 2025. doi: 10.1016/j.system.2024.103387

[39] C. Zhang, Y. Meng, and X. Ma. "Artificial intelligence in EFL speaking: Impact on enjoyment, anxiety, and willingness to communicate," *System*, vol. 121, 103259, 2024. doi: 10.1016/j.system.2024.103259

[40] V. Braun and V. Clarke, "Reflecting on reflexive thematic analysis," *Qualitative Research in Sport, Exercise and Health*, vol. 11, no. 4, pp. 589–597, 2019. doi: 10.1080/2159676X.2019.1628806

[41] D. Deckker and S. Sumanasekara, "Student-centric ethical frameworks for AI-driven education: Participatory consent and data ownership in a global perspective," *International Journal of Science and Research Archive*, vol. 15, no. 3, pp. 1491–1503, 2025. doi: 10.30574/ijsra.2025.15.3.1920