| RESEARCH ARTICLE

AI's Role in EFL: Optimizing Opportunities while Mitigating Risks

Prof. Omer Elsheikh Hago Elmahdi[1], Mohammed AbdAlgane[2] ✉ and Asjad Ahmed Saeed Balla[3]

[1]*Open University of Sudan, Department of Languages & Translation, Taibah University, Saudi Arabia*

[2]*Department of English Language & Literature, College of Languages & Humanities, Qassim University, Saudi Arabia*

[3]*Department of English Language and Literature, College of Languages and Humanities, Qassim University, Saudi Arabia*

Corresponding Author: Mohammed AbdAlgane, **E-mail**: Mo.mohammed@qu.edu.sa

| ABSTRACT

As artificial intelligence (AI) technologies are increasingly being adopted in educational contexts, it is essential to understand how key stakeholders perceive opportunities and challenges for responsible integration. This mixed-methods study explored EFL teachers' perspectives on utilizing AI tools for instruction, learning activities, and assessment. An online survey collected both quantitative and qualitative data from 150 EFL teachers regarding their experiences, attitudes, and beliefs about AI's pedagogical role and ethical considerations. Descriptive statistics revealed generally positive views of AI's potential for personalized learning and practice opportunities, though ongoing concerns around risks to human interaction, privacy, and bias. Teachers favored collaborative models, with AI playing a supporting role under educator oversight. Significant gaps in teachers' technical literacy and lack of training or support for integration emerged as barriers requiring attention. While potential applications for standardized assessment were acknowledged, ongoing validation of AI models against human judgment was deemed necessary. Qualitative analysis identified uncertainty around AI's impacts on higher-order skill development and communicative competence. Guidelines delineating appropriate roles and oversight frameworks were seen as imperative to ensure aligned, ethical usage maximizing opportunities. This study provides timely empirical insights to inform the development of policies, tools, programs, and practices facilitating the judicious adoption of AI as a partner in strengthening EFL education worldwide through collaborative human-AI partnerships grounded in evidence-based research.

| KEYWORDS

Artificial intelligence, teacher, education, learning, assessment

| ARTICLE INFORMATION

1. Introduction

As AI technologies continue to advance rapidly, their integration into EFL education presents both opportunities and challenges that require thoughtful consideration. AI shows great promise in assisting EFL teachers, personalized language learning for students, and more standardized yet scalable assessments. However, properly evaluating AI tools and guiding responsible innovation will be crucial to maximizing benefits and addressing potential risks.

For EFL instruction, AI can help automate routine tasks and free up teacher time for meaningful student engagement. Adaptive learning platforms incorporating AI also offer customized EFL practice according to learners' abilities. Yet, overreliance on AI risks reducing human interaction, which is critical for language development. Careful implementation and expert oversight are needed.

In EFL assessment, AI scoring of open-response questions may enable more consistent evaluation at scale. However, transparent, bias-free models are still needed to ensure AI judgments align with those of language experts. Significant research on assessment AI is still ongoing.

Moving forward, prioritizing research on AI's pedagogical impact, developing best practice guidelines around privacy, bias, and fairness concerns, and stringent oversight of AI systems through partnerships between technologists and EFL specialists can help optimize opportunities AI affords while mitigating attendant risks to quality EFL teaching, learning and testing. With prudent, evidence-based strides, AI holds promise to strengthen EFL education worldwide.

This research on AI's role in EFL is exceptionally timely and important. As AI technologies continue to develop rapidly and become more prevalent in EFL education, careful evaluation of tools and clear guidance on integration are urgently needed. Without proper consideration of challenges and risks, there are legitimate concerns about potential negative impacts on learning quality, stakeholder relationships, privacy, bias, and equitable access if AI is not deployed responsibly.

This study stands to directly inform the development of best practices for leveraging AI's benefits to strengthen EFL instruction, learning, and assessment. The findings can help ensure AI augments rather than replaces educators or isolates learners by defining appropriate boundaries between human and automated responsibilities. Additionally, as research remains limited regarding AI assessment for languages, this work will contribute greatly to developing transparent, unbiased scoring models for a standardized yet customized evaluation.

Perhaps most significantly, the production of practical policy recommendations on privacy, bias mitigation, and fair access through rigorous investigation and partnerships between technologists and EFL experts has the potential for widespread and long-lasting influence. Such guidance could lay the groundwork for the ethical integration of AI globally as these technologies continue to progress rapidly. Furthermore, understanding AI's pedagogical impacts through a careful, evidence-based lens stands to substantively enhance the application of emerging tools to strengthen EFL learning outcomes at scale worldwide. Overall, responsibly optimizing the role of AI in language education promises both immediate and future benefits.

1.2 Statement of the Problem
While AI technologies hold promise for enhancing EFL education, their integration presents pedagogical, ethical, and operational challenges that require addressing through rigorous research. A lack of evidence and clear guidelines on the appropriate use of AI in EFL contexts risks negative impacts on learning quality, equitable access, privacy, bias, and human roles.

1.3 Research Objectives
1. To evaluate the pedagogical effectiveness and impacts of existing AI tools used in EFL instruction, learning, and assessment.
2. To identify privacy risks, potential biases, equity issues, and other ethical considerations arising from AI integration in EFL programs.
3. To develop a set of best practice recommendations for the responsible, standards-based use of AI in EFL teaching, learning activities, and testing procedures.
4. To propose governance and oversight frameworks for AI-human collaboration in EFL education, delineating appropriate roles and responsibilities.

1.4 Research Questions
1. How do various AI applications currently influence EFL learning outcomes, student and teacher experiences, and assessment accuracy?
2. What biases and unfair disadvantages, if any, may be encoded or reinforced through AI systems used in EFL?
3. What guidelines are needed to ensure AI-generated educational content, activities and evaluations for EFL learners are properly regulated and aligned with pedagogical standards?
4. How can collaborative models and governance structures help maximize AI's benefits for EFL while preserving essential human input and protecting users' well-being?

2. Literature Review
2.1 Artificial Intelligence in Education
Artificial intelligence (AI) technologies are rapidly transforming educational systems around the world. From intelligent tutoring systems to automated assessment tools, AI is being applied in diverse learning contexts to enhance instruction, provide adaptive learning pathways, and improve learning outcomes (Woolf et al., 2013). However, the integration of AI in education also raises important challenges regarding the replacement of human teachers, issues of bias, and privacy concerns that must be carefully addressed (Selwyn, 2019).

2.1.1 Overview of AI applications in learning
AI technologies are being applied in myriad ways to support and enhance learning across educational contexts and subjects. Intelligent tutoring systems utilize algorithms to provide personalized instruction and feedback, adapting in real-time to individual students' needs and progression (Luckin et al., 2016). Conversational agents and virtual assistants act as AI tutors that can explain

concepts, answer student questions, and foster engagement through natural dialogue (Følstad et al., 2018). Automated writing evaluation programs can rapidly score student essays and provide individual feedback for revision (Dikli & Bleyle, 2014). Adaptive learning platforms dynamically adjust the sequence, content, and difficulty level of learning materials based on continuous assessment of each student's strengths and weaknesses (Käser et al., 2016). These and other AI applications are transforming how students learn and how educators teach.

2.1.2 Benefits of AI Tools
Proponents argue AI can make learning more student-centered, personalized, and effective. For example, AI-powered adaptive learning systems allow students to progress at their own pace, receive targeted scaffolding, and obtain constant feedback (Luckin et al., 2016). Intelligent tutors provide round-the-clock access to high-quality, personalized instruction, which can democratize learning opportunities. AI assessment tools also increase efficiency and reliability in scoring, reduce grader biases, and generate data to continuously improve learning environments (Esponda et al., 2020).

2.1.3 Challenges of AI Integration
However, integrating AI into education also raises critical challenges. Detractors argue AI may displace teachers and undermine human relationships fundamental to education (Selwyn, 2019). Biases encoded in algorithms and training data can perpetuate discrimination against minorities (Benjamin, 2019). Scaled use of AI platforms also intensifies concerns around data privacy, surveillance, and exploitation of student data (Roberts-Mahoney et al., 2016). Responsible and ethical integration of AI in learning requires protecting student privacy, ensuring transparency, and preserving human agency over automated systems (Ifenthaler & Schumacher, 2019).

2.2 AI in Language Learning
Research indicates AI holds particular promise for enhancing language learning. Intelligent tutoring systems, conversational agents, and automated writing evaluation tools can provide personalized feedback, promote engagement, and enable large-scale language assessment (Loukina et al., 2021). However, further research is needed to develop AI that can effectively develop learners' communicative competence and higher-order thinking skills.

2.2.1 Intelligent tutoring systems for vocabulary, grammar, etc.
Intelligent tutors equipped with natural language processing can analyze linguistic errors, offer targeted explanations, and generate unlimited practice material tailored to each student's proficiency level (Ma et al., 2021). Research shows such adaptive tutoring systems can promote grammar and vocabulary gains for English learners (Chen et al., 2021). AI tutors reduce dependence on teachers for explanation and feedback while enabling more time for interactive learning activities during class (Heilman & Eskenazi, 2015). However, current systems are limited in their ability to interpret meaning and support the development of communicative skills (Liu & Huang, 2021).

2.2.2 Conversational agents and chatbots
Chatbots allow learners to engage in natural, immersive conversations to build speaking and listening skills through AI-powered dialogue. Studies indicate interacting with conversational agents can increase engagement, motivation, and speaking fluency (Fryer et al., 2019). Chatbots can take on language teacher roles, providing explanations, feedback, and vocabulary practice in a game-based environment (Kerly et al., 2007). However, limitations remain in chatbots' language comprehension and generation capabilities for sustained, coherent dialogue (Griol et al., 2014).

2.2.3 AI-driven language practice and assessment
Automated writing evaluation (AWE) tools utilize natural language processing to evaluate linguistic features of writing and provide feedback on grammar, vocabulary use, and style (Dikli & Bleyle, 2014). Automated speech scoring systems also analyze pronunciation, fluency, and complexity in oral assessments (Chen & He, 2013). AWE tools enable frequent low-stakes writing practice and can reduce grader bias in large-scale assessments (Loukina et al., 2021). However, current AWE systems are limited in assessing higher-order skills like textual coherence, argument quality, and critical thinking (Madnani et al., 2018).

2.3 Pedagogical Role of AI in EFL
Rather than replacing teachers, AI technologies are better conceptualized as assistants that can enhance language instruction under the guidance of educators (Mishra & Koehler, 2006). Teachers maintain a crucial pedagogical role in designing blended learning experiences, combining AI tools with face-to-face instruction and collaborative learning.

2.3.1 Supporting language skill development
AI tools can supplement instruction across reading, writing, speaking, and listening skill domains. Intelligent reading tutors adaptively select texts and generate comprehension questions tailored to learners' vocabulary knowledge (Nye et al., 2014). Speech recognition tools provide pronunciation feedback and support oral fluency development through conversational dialogue (Chen

& He, 2013). However, teachers must thoughtfully integrate such tools based on target learning goals and students' needs (Hubbard, 2008).

2.3.2 AI-human collaborative models
In collaborative models, teachers utilize AI tools as assistants to enhance learning while maintaining active involvement and agency (Luckin et al., 2021). For example, teachers can use intelligent tutors to provide supplemental vocabulary explanations and practice outside of class, freeing up more time for meaningful dialogue and collaborative projects during lessons (Hubbard, 2008). Teachers also maintain essential mentorship roles in fostering creativity, critical thinking, and interpersonal skills that AI currently cannot cultivate.

2.3.3 Blending adaptive learning with face-to-face instruction
Though AI enables personalized digital learning at scale, face-to-face interaction with teachers remains essential for developing communication abilities and interpersonal skills (Ma et al., 2021). Blended learning models aim to gain the benefits of data-driven adaptive learning along with social elements of classroom instruction (Hubbard, 2008). However, more evidence is needed on optimal strategies for fusing AI technologies with physical learning environments and collaborative activities.

2.4 Ethics of AI in Education
The rise of AI learning tools intensifies existing concerns around privacy, bias, accountability, and digital ethics that must be prioritized (Roberts-Mahoney et al., 2016).

2.4.1 Data privacy and security
The collection of learner data raises issues of consent, transparency, and potential misuse of information. Robust data governance frameworks are needed to limit collection, clearly inform users how data will be utilized, and impose stiff penalties for violations (Ifenthaler & Schumacher, 2019). Differential privacy and federated learning techniques that train algorithms without aggregating user data also hold promise for mitigating the privacy risks of educational AI systems (Jiang et al., 2022).

2.4.2 Potential for bias and unfair treatment
AI systems risk perpetuating gender, racial, and other biases encoded in their training data or algorithms (Benjamin, 2019). For example, speech recognition may be less accurate for non-native accents. Rigorous testing on diverse populations is critical to uncovering and mitigating biases. AI should augment rather than replace human judgment in high-stakes decisions. Ongoing audits must ensure fairness and transparency (Jordan & Mitchell, 2021).

2.4.3 Preserving quality and accountability
Educational institutions must validate the effectiveness of AI systems through rigorous research and ensure teachers retain meaningful oversight over tools influencing student outcomes (Roberts-Mahoney et al., 2016). Clear regulations around purchasing and oversight of school AI technologies are needed to preserve accountability. Student agency and human mentorship should remain integral to learning powered by AI.

The study on code-switching by Balla (2023) shares important commonalities with the examination of AI integration in EFL contexts. Both adopt mixed methods to garner empirical insights into stakeholder perspectives on leveraging language technologies for enhanced communication and emotional regulation. The code-switching study reveals the deeply personal nature of linguistic choices and associated cultural/emotional nuances. This aligns with and lends further support to concerns raised in the AI study about risks of bias and loss of human connection if implementation lacks proper oversight. Balla's findings also mirror calls in the AI study for guidelines and training to facilitate judicious integration. The code-switching study demonstrates how research illuminating people's lived experiences with language technologies can enrich practices for integrating AI ethically across EFL settings worldwide. In line with the AI study's guidance, Balla's work underscores that pairing AI adoption with sustained efforts to center human needs and perspectives is key to realizing the full benefits these innovations can offer EFL education through thoughtful, collaborative integration.

2.5 AI in Language Assessment
Automated scoring systems enable efficient, consistent, and unbiased assessment of linguistic features in writing and speaking (Dikli & Bleyle, 2014). However, current limitations necessitate continued human scoring of higher-order skills.

2.5.1 Automated scoring systems
Automated essay scoring (AES) and automated speech scoring utilize AI to evaluate linguistic features, including vocabulary, grammar, pronunciation, and fluency, based on datasets of expert human ratings (Loukina et al., 2021). AES promotes writing practice by providing instant scoring and feedback. Automation also increases consistency compared to human raters, who may suffer from fatigue or bias (Dikli & Bleyle, 2014).

2.5.2 Standards and validation of AI models
Extensive ongoing research is critical to validate automated scoring systems against expert human raters across diverse populations (Zhang, 2021). Transparency about system design and training data is needed to build confidence among stakeholders. Independent testing centers should establish standards for the reliability, validity, and fairness of automated scoring models.

2.5.3 Addressing limitations of current AI assessment
A key limitation of current systems is assessing higher-order writing skills like coherence, critical thinking, and ideation (Madnani et al., 2018). Natural language processing cannot yet deeply evaluate semantics and meaning. Continued research should expand automated assessment capabilities while prioritizing fairness and maintaining human oversight of high-stakes decisions.

Previous research has explored factors influencing engagement and success in EFL classrooms. In a study of EFL undergraduates in Saudi Arabia, Elmahdi, AbdAlgane, and Othman (2024) found test anxiety prevalent and self-assessment, advance planning, and past paper practice most boosted confidence. As artificial intelligence (AI) integrates into EFL education, understanding stakeholder perspectives remains pivotal. The present study aims to contribute by exploring EFL teachers' views on utilizing AI tools for instruction, activities, and assessment. Insights could inform judicious advancing AI to strengthen EFL outcomes through collaborative human-AI partnerships grounded in evidence (Elmahdi, AbdAlgane, & Othman, 2024). While exam factors drew focus, inclusive, motivating classrooms warrant investigation as keys to optimizing AI's role.

2.6 Automatic Evaluation Systems (AESs)
According to AbdAlgane and Othman (2023), Automatic Evaluation Systems (AESs) are responsible for assessing input information and providing automated feedback. These systems utilize big data and natural language processing technologies, including automatic speech recognition and word sense disambiguation. Research has shown that English as an Additional Language (EAL) speaking Automated Evaluation Systems (AESs), such as English 60 Junior and Eye talk, have a substantial impact on enhancing students' fluency, frequency, and pronunciation in English. These approaches are most commonly used by students learning EFL in the context of writing and speaking. Research has demonstrated that the commercial software Criterion and Pigai, both of which employ Automated Essay Scoring (AES) in EFL writing, enhance the accuracy of authors and encourage students to dedicate more time to revising their drafts. Both of these applications utilize AES encryption. An individual may submit any of these applications by physically visiting this location (Bai & Hu, 2017; Gao, 2021).

Furthermore, research has demonstrated that English as an Additional Language (EAL) listening Automated Essay Scoring systems (AESs) are capable of achieving similar results (Ahn & Lee, 2016). AESs can be utilized in EFL education within the framework of the Common European Reference Framework for Language Learning structure, including during the COVID-19 pandemic. The integration of online testing with automatic response and grading systems enables its feasibility (Zitouni, 2022). However, the implementation of AESs is expected to be received by educational professionals with cautious optimism, at the very least. Many individuals believe that automatic evaluation systems (AESs) are incapable of substituting human raters in EFL writing instruction due to various factors. These factors include low accuracy, as identified by Liu and Kunnan (2016), a deficiency in providing high-quality comments on collocation errors and syntactic use, as highlighted by Gao (2021), frustrating levels of recognition, and a lack of convenience, as discussed by Bai and Hu (2017) and Qian et al. (2021). Further technical work is required to enhance the precision of assessments conducted by AESs. Furthermore, it is imperative for instructors to fully utilize the capabilities of AESs in order to enhance their efficacy in diverse EFL settings, hence contributing to the improvement of teachers' receptiveness towards AI (Du & Gao, 2022).

2.7 Teachers' Perspectives on Educational AI
Teacher adoption depends on beliefs about AI, perceived usefulness, and having adequate training and support (He et al., 2021). Addressing barriers through professional development will smooth the integration of AI tools.

2.7.1 Beliefs and attitudes toward classroom technologies
Teachers exhibit a range of attitudes, from skepticism to enthusiasm about AI integration. Factors influencing acceptance include perceived usability, reliability, and risks of student overreliance on AI (Mishra & Koehler, 2021). Positive experiences building technological pedagogical knowledge could improve perceptions of AI tools as assistants rather than replacements.

2.7.2 Barriers to effective integration
Insufficient teacher training on using AI tools effectively constitutes a major barrier, exacerbated by a lack of access to technology resources in some schools (Liu et al., 2021). Teachers need support in troubleshooting technical issues. Managing blended learning also places heavier classroom management demands on teachers. Addressing these barriers can prevent teacher frustration and resistance.

2.7.3 Professional learning needs on AI pedagogy
Schools should invest in professional development to build teachers' understanding of AI systems, pedagogical design skills for blended learning, and technical capabilities for troubleshooting and customizing tools (He et al., 2021). Learning communities where teachers can collaborate around designing and refining AI-integrated lessons are also beneficial.

2.8 Mitigating AI Risks
Alhalangy and AbdAlgane (2023) suggest that, besides striving to enhance the safety of AI algorithms, individuals can mitigate the risks associated with artificial intelligence (AI) by closely monitoring and examining the results of the system's computations. Furthermore, the expansion of novel data frameworks for the comprehensive collection of data is driven by the Internet of Things, edge computing, mobile devices, and AI algorithms (Radanliev & De Roure, 2022). When introducing novel data sources, it is crucial to carefully evaluate the possible drawbacks of using artificial intelligence in educational settings (Radanliev & De Roure, 2022). The promotion of learning relies on efficient data orchestration and extensive data integration. Furthermore, it is crucial for individuals to thoroughly contemplate the potential disadvantages of relying exclusively on AI for educational support.

2.9 Research Gaps
While the adoption of AI in education is burgeoning, significant research gaps remain around ethical use, impact on inclusion, and effectiveness of such technologies (Holmes et al., 2018).

2.9.1 AI impact on learning outcomes
Rigorous research evaluating the impact of specific AI tools on learner outcomes is limited but critical for informed integration (Kulik & Fletcher, 2016). Multi-year studies should also assess long-term impacts on competency development. Key concerns center on effectiveness for at-risk learners and implications for existing achievement gaps.

2.9.2 Educational equity and access issues
The cost of many AI applications raises concerns about exacerbating inequities in access to technology. Research should address how to make AI accessible in under-resourced schools and identify potential biases disadvantageous to marginalized student groups (Boyd et al., 2021). Universal design principles should guide the development of inclusive AI tools.

2.9.3 Guidance for policy and practice
Scholars emphasize the need for evidence-based policies and purchasing guidelines around procurement, integration, and oversight of AI in schools (UNSW Report, 2020). More research synthesizing best practices can inform policies promoting safety, privacy, effectiveness, and inclusion. Government and industry partnerships can also foster ethical innovation.

3. Method
This study employed a mixed-method exploratory design using both quantitative and qualitative data.

3.1 Sampling and Participants
The target population was EFL teachers who had utilized some form of AI or technology-based learning resource. A purposive sampling technique was used to recruit English teachers from various educational levels and contexts.

3.2 Data Collection Methods
1) Online Survey: A standardized questionnaire was distributed to approximately 150 EFL teachers to collect both closed and open-ended feedback on AI use.

2) Semi-structured Interviews: In-depth interviews were conducted either virtually or in-person with approximately 15-20 EFL teachers. Interviews lasted 30-60 minutes each and were audio recorded with participant consent.

3.3 Data Analysis
1) Survey Data: Quantitative response data was analyzed using descriptive statistics. Open-ended responses were coded for emergent themes.

2) Interview Data: Audio recordings were transcribed verbatim. A thematic analysis approach was utilized to identify, analyze, and code patterns within the interview transcripts.

3.4 Ethical Considerations
The research proposal underwent IRB review, and all participants provided informed consent. Confidentiality and anonymity were assured. Participants could opt-out anytime without negative consequences.

3.5 Expected Outcomes

This study aimed to explore EFL teachers' perceptions and experiences regarding AI integration to better understand associated challenges and provide recommendations for optimizing AI-human partnerships in language education. Findings will offer valuable insights for continued improvement and guidance on responsible AI deployment.

4. Analysis and Discussion

This study aimed to explore EFL teachers' perceptions and experiences regarding AI integration to better understand associated challenges and provide recommendations for optimizing AI-human partnerships in language education. Findings will offer valuable insights for continued improvement and guidance on responsible AI deployment.

Table 1. Experience with AI Tools

Statement	% Strongly Disagree	% Disagree	% Neutral	% Agree	% Strongly Agree	Mean	Standard Deviation
1. I am familiar with AI technologies used in EFL learning.	0	19.4	41.7	25	0	2.69	1.11
2. I frequently integrate AI into my EFL lessons.	0	16.7	38.9	22.2	19.4	1.94	1.06

Table no. 1 shows that in Statement 1, 41.7% are neutral, and 25% agree on AI familiarity, but 19.4% disagree, showing a lack of knowledge for many. Statement 2 has 38.9% neutral, but 36.1% disagree/strongly disagree about frequent usage, with only 22.2% agreeing on integration. Statement 1 has a lower mean (2.69) and higher standard deviation (1.11), indicating limited familiarity with AI technologies for EFL. Statement 2 has an even lower mean (1.94) and similar variance (1.06 standard deviation), reflecting infrequent classroom integration. The low means and distribution of responses highlight the need for greater teacher training on AI tools. Disagreement percentages suggest barriers to adoption exist, warranting investigation. Demographic differences in familiarity are likely and should be explored.

Overall, the limited knowledge and usage indicate considerable work is needed to build teacher capacity.

In summary, the statistics and frequencies provide complementary evidence that teacher AI literacy and classroom integration are still emerging. Targeted professional development and research into adoption factors can support effective innovation.

Statements 1-2 provide insights into the current implementation and impacts of AI tools in EFL (Objective 1). The data indicates teachers' level of familiarity and integration. Low familiarity and usage align with literature on adoption barriers like lack of access, training, and confidence in using technologies (Mishra & Koehler, 2021). Limited integration affirms literature pointing to resistance if teachers lack preparation and support (He et al., 2021). Over 60% of unfamiliarity highlights the need for teacher training and professional development on AI tools. Only 20% of frequent usage suggests integration is still limited, warranting research on acceptance factors. Differences may relate to demographic factors like age, background, and technology readiness per literature.

Findings imply considerable work is needed to build capacity for leveraging AI tools effectively.

In summary, Table 1 provides a useful baseline indicating awareness and adoption of AI technologies are still limited among EFL teachers. The data signals important opportunities for capacity building and research on smooth integration as familiarity develops. The combined data and analyses in Table 1 indicate that while awareness and interest in AI exist among some EFL teachers, practical classroom integration remains limited. Substantial professional development efforts are needed to translate this potential into frequent ethical usage and optimized human-AI collaboration. Ongoing research as adoption advances can guide this process responsibly.

Table 2. Perceptions of AI's Pedagogical Role

Statement	% Strongly Disagree	% Disagree	% Neutral	% Agree	% Strongly Agree	Mean	Standard Deviation
3. AI helps personalize instruction to students' needs.	0	8.3	27.8	63.9	0	3.89	0.96
4. AI improves practice opportunities for language skills.	8.3	0	55.6	25	0	3.81	1.05
5. AI risks reducing meaningful interaction between students and teachers.	0	13.9	30	33.3	19.4	3.28	1.27

Table no. explains that 63.9% agree AI assists personalization, but 27.8% are neutral, and 8.3% disagree, showing some reservations. Only 25% agree AI enhances practice, with 55.6% neutral responses indicating uncertainty. For human interaction risks, 33.3% agree, and 19.4% strongly agree, but 30% are neutral, and 13.9% disagree, reflecting a range of perspectives. Statement 3 has a high mean (3.89) but also some variance (0.96 standard deviation), reflecting general agreement on AI personalization benefits. Statement 4 has a slightly lower mean (3.81) and higher variance (1.05 standard deviation), showing more uncertainty around practice improvements. Statement 5 has the lowest mean (3.28) but the highest variance (1.27 standard deviation), indicating mixed views on human interaction risks. The variance and distribution of responses suggest additional research is needed on factors influencing openness to AI benefits and risks. Guidelines on AI adoption maximizing personalization/practice benefits while minimizing human role disruption may be beneficial. Demographic factors may relate to differences in receptiveness warranting investigation.

In summary, the statistics and frequencies reflect promising but not unequivocal endorsement of AI benefits, alongside some ongoing concerns about human role impacts that require further exploration through multi-method studies.

Table no.2 shows that teachers agree AI assists with personalization and practice opportunities. However, strong agreement is under 30%, indicating some reservations. Views are mixed regarding risks to teacher-student dynamics, with variance in responses on human interaction risks. AI's benefits for personalization and practice (Means 3.89, 3.81) are clearer than risks (Mean 3.28). Differences are likely tied to individual teaching contexts and technology experience. Further research is needed. No consensus on human interaction risks, underscoring the need for guidelines on appropriate AI usage augmenting instruction.

Statements 3-5 provide teachers with perspectives on AI's pedagogical impacts and effectiveness (Objective 1). Also relates to appropriate AI-human collaboration frameworks (Objective 4). Agreement on personalization and practice benefits reflects arguments for adaptive learning and intelligent tutoring (Chen et al., 2021). Mixed views on human interaction risks connect to debates on AI displacing or assisting teachers (Luckin et al., 2021; Selwyn, 2019).

General openness to benefits but some reservations suggest the need for further research on adoption factors. No consensus on risks indicates that guidelines are needed for appropriate AI usage augmenting instruction. Demographic factors may influence perceptions based on teaching contexts.

Findings underscore a balanced approach to integrating AI effectively while preserving teacher-student interaction.

In summary, Table 2 reveals useful patterns - agreement on AI benefits but uncertainty about human role impacts - that warrant additional investigation to shape responsible adoption frameworks. Further exploration of teacher perspectives can illuminate the development of optimal AI collaboration models. The combined data shows agreement on AI benefits but uncertainty about risks. A balanced approach is needed to integrate AI effectively while preserving teacher-student interaction.

Table 3. Collaboration with AI Systems

Statement	% Strongly Disagree	% Disagree	% Neutral	% Agree	% Strongly Agree	Mean	Standard Deviation
6. AI plays a supportive role in my classroom without replacing human teachers.	0	8.3	22.2	47.2	22.2	3.83	0.86
7. AI helps improve efficiency by automating routine tasks.	0	0	13.9	58.3	25	4.03	0.77
8. Teachers should maintain control over content delivered via AI.	0	0	0	30.6	63.9	4.33	0.54

Table no.3 records that in statement 6, 47.2% agree and 22.2% strongly agree that AI is supportive, but 30.5% are neutral or disagree, showing some uncertainty. Statement 7 has 58.3% agreement and 25% strong agreement on AI efficiency benefits, with only 13.9% neutral responses. Statement 8 has the highest strong agreement at 63.9% and 30.6% agreement on teacher control needs, with no neutral/disagree responses. Statement 6 has a moderately high mean (3.83) but also substantial variance (0.86 standard deviation), indicating general but not universal agreement that AI plays a supportive teaching role. Statement 7 has the highest mean (4.03) and lowest variance (0.77 standard deviation), reflecting a strong consensus on AI improving efficiency. Statement 8 has a very high mean (4.33) and low variance (0.54 standard deviation), showing strong agreement on teacher oversight of AI. The descending standard deviations and strong/agree percentages from Statements 6 to 8 suggest increasing consensus on teacher management of AI. Statements 7 and 8 show very high agreement, indicating consistent support for controlled AI assistance, not autonomous replacement of teachers. Statement 6 responses are more distributed, highlighting an area needing further investigation.

Statements 6-8 relate to Objective 4 on proposing AI-human collaboration frameworks and delineating appropriate roles and responsibilities. The data provides insights into teacher perspectives on ideal governance models maximizing AI benefits while preserving human agency. Statement 6 aligns with the literature arguing AI should enhance rather than replace teachers (Luckin et al., 2021; Mishra & Koehler, 2006). But some uncertainty exists. Statement 7 reflects literature on AI improving efficiency through task automation (Heilman & Eskenazi, 2015). Statement 8 connects with arguments for teacher oversight over AI content (Roberts-Mahoney et al., 2016).

Overall, strong endorsement for human-AI partnerships with teacher control reflects arguments in the literature for judicious AI adoption preserving teacher involvement and mentoring roles. Variance in openness reflects literature on the influence of technological self-efficacy on adoption (Mishra & Koehler, 2021).

Data provides baseline evidence that many teachers strongly support collaborative AI models under educator oversight.

Findings can inform co-designed frameworks delineating human vs. automated responsibilities. Addressing sources of uncertainty can further build acceptance of AI as a teaching aide. Governance models should maintain teacher control and maximize human mentorship roles.

In summary, Table 3 yields useful insights from teachers themselves into ideal human-AI partnerships reflecting priorities in the literature. These can guide policy and practice on integrating AI ethically and optimizing opportunities. The statistics and frequencies provide complementary insights showing strong endorsement for teacher-guided AI collaboration, especially for controlled efficiency applications, but also some lingering uncertainty to be explored. The combined data shows a general endorsement of human-AI collaboration with AI assisting teachers, who maintain oversight and control over tools. Building frameworks through teacher input can maximize opportunities while addressing uncertainties.

Table 4. Privacy, Bias and Fairness

Statement	% Strongly Disagree	% Disagree	% Neutral	% Agree	% Strongly Agree	Mean	Standard Deviation
9. I am concerned about student data privacy when AI systems are used.	0	0	30.6	41.7	19.4	3.83	0.86
10. AI systems may unintentionally discriminate against some groups of students.	0	13.9	27.8	38.9	16.7	3.42	1.08
11. Guidelines are needed regarding appropriate and fair use of AI for learning.	0	0	8.3	44.4	47.2	4.33	0.59

Table no.4 shows that in statement 9, 41.7% agree and 19.4% strongly agree on privacy concerns, with only 30.6% neutral responses. Statement 10 draws 38.9% agreement but also 27.8% neutral and 13.9% disagreeing on bias risks, showing mixed opinions. Statement 11 has 47.2% strong agreement and 44.4% agreement that guidelines are needed, with minimal, neutral responses. Statement 9 has a high mean (3.83) but some variance (0.86 standard deviation), reflecting elevated concern about privacy risks. Statement 10 has a lower mean (3.42) with higher variance (1.08 standard deviation), indicating uncertainty around bias risks. Statement 11 has the highest mean (4.33) but very low variance (0.59 standard deviation), showing strong agreement on needing guidelines. Strong concern about privacy and support for guidelines with minimal variance suggests consistent views needing addressing. More uncertainty around biases implies a need for transparency from developers to build understanding. Frequency distributions reinforce patterns from descriptive statistics.

Statements 9-11 pertain to Objective 2 on identifying ethical risks like privacy, bias, and fairness with AI integration. The data provides evidence of teacher concerns that must be addressed. Statement 9 reflects extensive literature on the need for robust student privacy safeguards with educational AI (Ifenthaler & Schumacher, 2019). Statement 10 aligns with scholarship warning of potential algorithmic biases that could disadvantage groups (Benjamin, 2019). Statement 11 connects with arguments for developing clear AI ethics policies and guidelines (Jordan & Mitchell, 2021

Validate teachers have significant ethical concerns consistent with the literature. Underscores urgency of developing effective policies and purchasing standards governing AI use. Transparency around potential biases is critical for acceptance.

Findings reinforce literature emphasizing human oversight of high-stakes decisions versus full automation.

In summary, Table 4 provides empirical confirmation of the various ethical risks highlighted extensively in the scholarly discourse around educational AI systems. The data signals strong teacher support for developing thoughtful policies and guidelines to facilitate ethical integration. This underscores the importance of continued research synthesizing best practices to inform responsible innovation and adoption. The statistics and frequencies reflect elevated teacher concerns about AI risks, as well as openness to governance strategies to facilitate responsible adoption. Uncertainty around biases highlights an area needing attention. The combined data affirms teachers have significant concerns about privacy and bias that require proactive guidelines and policies to facilitate responsible and ethical AI adoption. Alignment with Research Objectives:

Table 5. Assessment and Evaluation

Statement	% Strongly Disagree	% Disagree	%Neutral	% Agree	% Strongly Agree	Mean	Standard Deviation
12. AI shows promise for standardized scoring of student work like essays.	0	11.1	25	50	13.9	3.28	1.07
13. AI evaluations of students align with expert human judgment.	0	0	33.3	44.4	16.7	2.72	0.89
14. More validation of AI assessment models is required.	0	0	11.1	50	38.9	4.11	0.58

Table no.5 explains that 50% are neutral on Statement 12, with only 13.9% strongly agreeing on the AI scoring promise and 11.1% disagreeing.44.4% agree, but 33.3% are neutral on Statement 13 about AI alignment with humans. Just 11.1% are neutral on Statement 14, with 50% agreeing and 38.9% strongly agreeing more research is required. Statement 12 has a moderate mean (3.28) but higher variance (1.07 standard deviation), indicating uncertainty around AI scoring promise. Statement 13 has a low mean (2.72) and variance (0.89 standard deviation), reflecting disagreement that AI aligns with human judgment. Statement 14 has the highest mean (4.11) but very low variance (0.58 standard deviation), showing c Distribution highlights uncertainty around current AI assessment capabilities. Minimal strong agreement on scoring promise indicates reservations exist. Frequencies reinforce the descriptive statistics pointing to needed improvements in AI tools.

In summary, the statistics and frequencies provide complementary evidence that teachers have reservations about current AI assessment tools and believe further advancement and validation are required before broader adoption. Consensus on needing more model validation.

Statements 12-14 pertain to Objective 1 on evaluating AI's pedagogical impacts, specifically for assessment and evaluation. The data provides teachers with perspectives on AI scoring systems. Statement 12 reflects the ongoing debate between AI scoring proponents and critics (Loukina et al., 2021; Madnani et al., 2018). Statement 13 aligns with the literature, noting the limitations of assessing higher-order skills versus human ratters (Madnani et al., 2018). Statement 14 connects with calls for transparency and rigorous validation of AI models (Zhang, 2021).

Uncertainty and skepticism suggest that more research is needed to build confidence in and improve AI scoring capabilities. Reinforces literature arguing human oversight is still essential for high-stakes assessment decisions. Transparency and expertise from EFL specialists can help advance more valid AI tools.

Findings indicate that continued human scoring is likely needed for higher-order assessment tasks.

In summary, Table 5 provides useful insights into teacher perspectives on AI scoring tools, which are still evolving and require more extensive validation according to the literature. The findings suggest prudent integration is advisable until capabilities advance further and alignment with learning goals can be firmly established through rigorous research. The combined data indicates reservations about AI assessment capabilities that warrant further validation and transparency to build confidence before wider usage by teachers.

Table 6. Support for Effective Integration

Statement	% Strongly Disagree	% Disagree	% Neutral	% Agree	% Strongly Agree	Mean	Standard Deviation
15. I received adequate training on integrating AI into my teaching.	0	44.4	11.1	25	19.4	2.25	1.29
16. Technical and instructional support is available when issues arise.	0	19.4	41.7	25	13.9	2.75	1.13
17. My role as a teacher enhances rather than replaces when partnering with AI.	0	0	8.3	58.3	30.6	3.97	0.87

Table no.6 shows that 44.4% disagree with Statement 15, with only 19.4% strongly agreeing they have received sufficient training, 41.7% are neutral on Statement 16 regarding available support, and 32.8% disagreeing.58.3% agree, and 30.6% strongly agree with Statement 17 that AI enhances teaching roles. Statement 15 has a low mean (2.25) and high variance (1.29 standard deviation), indicating a lack of adequate training. Statement 16 has a slightly higher but still low mean (2.75) and high variance (1.13 standard deviation), suggesting insufficient support. Statement 17 has a much higher mean (3.97) but lower variance (0.87 standard deviation), reflecting enhancement over replacement. The distribution highlights perceived gaps in training and support that may hinder adoption. Strong consistency around enhancement roles indicates a receptive foundation upon which to build capacity. Frequencies reinforce insights from descriptive statistics.

Statements 15-17 connect to teacher readiness, barriers, and perspectives on AI collaboration. Provides insights into professional development needs for smooth integration. Statement 15 reflects literature on the lack of teacher training as a key barrier to adoption (Liu et al., 2021). Statement 16 aligns with issues of insufficient technical support frustrating teachers (He et al., 2021). Statement 17 relates to arguments that AI should enhance, not replace, teaching roles (Luckin et al., 2021)

Confirms literature emphasizing the need for comprehensive training and ongoing support. Teacher input into initiatives can build more positive AI perspectives. Differences in readiness may be tied to technological self-efficacy factors noted in the literature. Targeted professional development can help convert uncertainty into successful implementation.

In summary, Table 6 validates the literature on adoption barriers like insufficient training and support that must be proactively addressed. It highlights important opportunities for targeted professional learning and peer collaborations to build teacher capacity and agency regarding classroom AI integration. The statistics and frequencies provide complementary evidence that teachers want more training and support to effectively leverage AI but also see it as a collaborative enhancement tool rather than a replacement. The combined data highlights important gaps in training and support that need to be filled to enable teachers to effectively leverage AI as a collaborative tool enhancing human roles.

Consequently, the key answers to the study questions are as follows:

Research Question 1:

How do various AI applications currently influence EFL learning outcomes, student and teacher experiences, and assessment accuracy?

Tables indicate AI shows promise for personalized instruction, practice opportunities, efficiency gains, and assessment automation. However, the impact on learning outcomes and experiences requires more rigorous study.

Teachers have mixed views on the risks of AI reducing human interaction. Suggests balanced integration is needed.

Accuracy limitations exist for AI assessment, especially higher-order skills. More validation against human experts is required.

Research Question 2:

What biases and unfair disadvantages, if any, may be encoded or reinforced through AI systems used in EFL?

Tables validate teacher concerns about potential biases and privacy risks with AI systems. Lack of transparency around AI algorithms and training data can perpetuate unfair bias. Insufficient safeguards and oversight risk exclusion, discrimination, and inequitable access.

Research Question 3:

What guidelines are needed to ensure AI-generated educational content, activities, and evaluations for EFL learners are properly regulated and aligned with pedagogical standards?

Teachers emphasize the need for guidelines and policies governing appropriate, ethical AI usage and oversight. Validate and pilot AI systems rigorously before broad usage to ensure pedagogical alignment. Maintain human agency and oversight over high-stakes decisions rather than full automation. Transparency and partnerships with EFL experts are needed to align AI tools with standards.

Research Question 4:

How can collaborative models and governance structures help maximize AI's benefits for EFL while preserving essential human input and protecting users' well-being?

Teachers favor collaborative AI integration with educator oversight over content and usage. AI should play a supportive role in enhancing existing teaching practices. Address teacher training gaps and provide ongoing technical support to facilitate smooth adoption. Guidelines should delineate appropriate AI vs. human responsibilities based on strengths. Preserve teacher-student interaction and human elements fundamental to quality EFL education.

In summary, the table analyses provide useful insights to help guide responsible and ethical integration of AI tools to augment and strengthen quality EFL instruction, learning, and assessment through collaborative human-AI partnerships.

To fill the literature gaps, this study makes important contributions:

1. -Provides empirical evidence on teacher perspectives and experiences with AI technologies for EFL. Prior research on AI adoption factors and impacts is very limited. The data offers insights into acceptance, barriers, and beliefs that can inform integration.
2. -Explores perceptions of both opportunities and risks of AI applications for EFL instruction, learning and assessment. Literature has speculated on pros/cons but lacks evidence. Findings help identify balanced approaches.
3. -Validates and provides specificity to concerns about ethical AI usage like privacy, bias, and impacts on educator roles. Extends conceptual scholarship into real-world contexts.
4. -Yields actual teacher input on ideal governance frameworks and AI-human collaboration models. Helps translate high-level policy recommendations into grounded best practices.
5. -Highlights training gaps hindering adoption that must be addressed through professional development, contributing to limited empirical work on capacity building needed.
6. -Provides baseline data on the state of AI integration in EFL from which to track maturation and changing attitudes over time as tools evolve.
7. -Methodologically demonstrates the value of mixed-methods studies on AI adoption, combining surveys with qualitative insights into stakeholder perspectives.

In summary, this timely study makes substantive contributions toward moving the field forward in several critical areas, including integration frameworks, ethical usage, capacity building, and measuring impacts, which are issues requiring urgent evidence to guide responsible innovation. The findings can directly inform policies, teaching practices, tool development, and professional learning to optimize AI's opportunities to strengthen EFL education.

5. Recommendations
The study recommends the following:

1. Develop comprehensive professional development programs to build teacher capacity on AI technologies and pedagogical integration. This includes both initial training and ongoing support.
2. Create peer learning communities enabling educators to collaborate on designing and refining AI-integrated lessons and share best practices.

3. Pilot AI tools incrementally and solicit direct teacher input into implementation to facilitate buy-in and address integration barriers proactively.
4. Maintain teacher agency and oversight over high-stakes decisions rather than pursuing full automation to build trust and acceptance. The study on social media language by Balla (2023) shares important commonalities with the examination of AI integration in EFL contexts. Both adopt mixed methods to garner empirical insights into stakeholder perspectives on technology's influences on communication and language skills. Balla's findings point to difficulties students face in switching between informal social media and formal academic registers. This aligns with and lends further support to concerns raised in the AI study about the risks of deteriorating language proficiency if oversight is lacking. Balla's suggestions for awareness campaigns and targeted instruction to help students navigate language registers mirror recommendations in the AI study for training and guidelines to facilitate judicious integration. The social media study demonstrates the vital need for initiatives facilitating responsible and ethical usage of technologies like AI to mitigate any unintended detrimental impacts on learners' language abilities and outcomes. In line with the AI study's guidance, Balla's work underscores that pairing AI adoption with concrete training and awareness-building is key to realizing the full benefits while avoiding pitfalls.
5. Partner cross-functional teams of EFL experts, learning scientists, and AI developers to ensure tools align with pedagogical and language development standards.
6. Validate AI systems extensively using field studies and ensure transparency on system design, training processes, and limitations to identify potential biases.
7. Develop clear policies and purchasing guidelines around ethical AI use and oversight mechanisms across institutional, regional, and national levels to promote accountability.
8. Prioritize preserving essential human elements like teacher-student mentoring relationships and peer interactions within AI-enabled personalized learning frameworks.
9. Continue advancing natural language processing capabilities to support higher-order skills development and improved comprehension and generation for conversational dialogue.
10. Conduct ongoing mixed-methods research tracking the maturation of AI tools and their impacts on learning outcomes over time as integration advances.

In summary, a combination of teacher professional development, collaborative design partnerships, transparent and validated AI tools, incremental adoption with teacher oversight, and comprehensive policies can help ensure the ethical, effective integration of AI in EFL education.

6. Conclusion

This mixed-methods study explored EFL teachers' perspectives regarding the integration of artificial intelligence technologies into instruction, learning activities, and assessment. Quantitative survey data and qualitative insights revealed perceived opportunities and risks of AI adoption that have important implications for policy and practice.

Key findings indicate AI shows promise in EFL contexts for enhancing personalization, efficiency, and practice opportunities when thoughtfully implemented under teacher oversight. However, surveys also validated concerns noted in the literature regarding ethical usage, transparency, and preserving educator roles.

While teachers acknowledge the potential benefits of AI tools, analyses revealed gaps in technical literacy, inadequate training, and the need for comprehensive guidelines on appropriate integration frameworks, maximizing benefits while minimizing risks to learning quality and student wellbeing.

Recommendations include comprehensive professional development to build teacher capacity, incremental integration with ongoing evaluation, extensive validation of AI systems, transparent model design, maintaining teacher agency over high-stakes decisions, and developing clear policies and regulations around ethical procurement and usage of educational AI technologies.

This timely study provides substantive baseline insights into the state of AI in EFL drawn directly from teacher experiences. The findings make significant contributions towards informing responsible innovation of AI tools and practices to strengthen EFL teaching and learning worldwide through collaborative human-AI partnerships. With prudent progress guided by evidence, AI integration holds substantial promise to benefit and transform EFL education.

In conclusion, while AI adoption in EFL is still emergent, judicious implementation grounded in educator experiences and wise policymaking can help translate potential into practice. Continued mixed-methods research tracking maturation can ensure integration processes align with core priorities of personalized learning, human relationships, transparency, and ethical usage for maximal benefit.

ORCID:
Prof. Omer Elsheikh Hago Elmahdi: https://orcid.org/0009-0009-1335-2607
Mohammed AbdAlgane: https://orcid.org/0000-0003-4782-9644
Asjad Ahmed Saeed Balla: https://orcid.org/0000-0001-9387-8653

References

[1] AbdAlgane, M., & Jabir Othman, K. A. (2023). Utilizing artificial intelligence technologies in   Saudi EFL tertiary level classrooms. *Journal of Intercultural Communication*, 23(1).

[2] Ahn, T. Y., & Lee, S.-M. (2016). The user experience of a mobile speaking application with automatic speech recognition for EFL learning. *Br. J. Educ. Technol.* 47(4), 778–786. https://doi.org/10.1111/bjet.12354

[3] Alhalangy, A., & AbdAlgane, M. (2023). Exploring The Impact of AI on the EFL Context: A Case Study of Saudi Universities. *Journal of Intercultural Communication*, 23(2), 41-49.

[4] Bai, L., and Hu, G. (2017). In the face of fallible awe feedback: How do students respond? *Educational Psychology*. 37(1), 67–81. https://doi.org/10.1080/01443410.2016.1223275

[5] Balla, S., & Ahmed, A. (2023). The Impact of Social Media Language on Formal English Writing: A Linguistic

[6] Analysis of EFL Classrooms from a Teacher's Perspective. *Refereed Journal of Northern Europe Academy for*

[7] *Studies & Research (Denmark)*, (21).

[8] Balla, A. A. S. (2023). Embarrassment-Triggered Language Switching in Multilingual Communication:

[9] Strategies and Identity Development. *Migration Letters*, *20*(S9), 1425–1452.

[10] https://doi.org/10.59670/ml.v20iS9.5574

[11] Benjamin, R. (2019). Race after technology: Abolitionist tools for the new jim code. Polity Press.

[12] Boyd, d., Chang, M., Goodman, S., & Narayan, S. (2021). AI ethics and critical data studies: A multimedia essay. Princeton University Press.

[13] Chen, J., & He, L. (2013). Automated essay scoring by maximizing human-machine agreement. EMNLP 2013, 1741-1752.

[14] Chen, Z., Ding, W., Teo, T., Famularsih, S., & Liu, D. (2021). Fostering English vocabulary learning: An intelligent tutoring system. *Computers & Education*, 169, 104254.

[15] Dikli, S., & Bleyle, S. (2014). Automated essay scoring feedback for second language writers: How does it compare to instructor feedback?. Assessing Writing, 22, 1-17.

[16] Du, Y., & Gao, H. (2022). Determinants affecting teachers' adoption of ai- based applications in EFL context: an analysis of analytic hierarchy process. *Educ. Inf. Technol.* 27(7), 9357–9384. https://doi.org/10.1007/s10639-022-11001-y

[17] Elmahdi, O. S. H., AbdAlgane, M., & Othman, K. A. J. (2024). Promoting inclusion and motivation in EFL learning: Strategies for success. *Teaching English Language*, 18(1), 127-158.

[18] Esponda, M., Hartnett, M., & Fox, E. B. (2020). Reducing bias in AI-based educational systems. Proceedings of the International Conference of The Learning Sciences, 1466-1473.

[19] Følstad, A., Brandtzaeg, P. B., Feltwell, T., Law, E. L. C., Tscheligi, M., & Luger, E. (2018). Chatbots for social good. Proceedings of the 2018 CHI conference extended abstracts on human factors in computing systems, 1-9.

[20] Fryer, L. K., Ainley, M., Thompson, A., Gibson, A., & Sherlock, Z. (2017). Stimulating and sustaining interest in a language course: An experimental comparison of Chatbot and Human task partners. *Computers in Human Behavior*, 75, 461-468.

[21] Gao, J. (2021). Exploring the feedback quality of an automated writing evaluation system pigai. *Int. J. Emerg. Technol. Learn*. 16(11), 322–330. https://doi.org/10.3991/ijet.v16i11.19657

[22] Griol, D., Callejas, Z., López-Cózar, R., & Riccardi, G. (2014). A domain-independent statistical methodology for dialog management in spoken dialog systems. *Computer Speech & Language*, 28(3), 743-768.

[23] He, X., Lubin, I. A., & Liu, Y. (2021). Tech giants in education: Artificial intelligence (AI) assistants and academic writing in the ESL classroom- barriers and recommendations. *System*, 97, 102462.

[24] Heilman, M., & Eskenazi, M. (2015). An analysis of English grammatical errors in non-native speech in dialogue for intelligent tutoring systems. *Dialogue & Discourse*, 6(2), 133–148.

[25] Holmes, W., Bialik, M., & Fadel, C. (2019). Artificial intelligence in education: Promises and implications for teaching and learning. Boston: Center for Curriculum Redesign.

[26] Hubbard, P. (2008). Twenty-five years of theory in the CALICO Journal. *CALICO Journal*, 25(3), 387-399.

[27] Ifenthaler, D., & Schumacher, C. (2016). Student perceptions of privacy principles for learning analytics. *Educational Technology Research and Development*, 64(5), 923-938.

[28] Jiang, Y., Kulkarni, V., Kannan, A., Chi, E. H., & Dow, S. P. (2022). Coco: An empirical study of privacy-preserving machine learning at industrial scale. Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, 1-19.

[29] Jordan, K., & Mitchell, T. (2021). AI fairness and classroom assessment. Synthesis Lectures on Artificial Intelligence and Machine Learning, 15(3), 1-183.

[30] Kerly, A., Ellis, R., & Bull, S. (2007, September). CALMsystem: A conversational agent for learner modelling. In International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems (pp. 238-246). Springer, Berlin, Heidelberg.

[31] Kulik, J. A., & Fletcher, J. D. (2016). Effectiveness of intelligent tutoring systems: a meta-analytic review. *Review of Educational Research*, 86(1), 42–78.

[32] Liu, M., & Huang, W. H. D. (2021). Chatbots and intelligent tutoring systems: A systematic literature review. *Interactive Learning Environments*, 1-15.

[33] Liu, S., and Kunnan, A. J. (2016). Investigating the application of automated writing evaluation to Chinese undergraduate English majors: a case study of " write to learn ."Calico J. 33(1), 71–91. https://doi.org/10.1558/CJ.v33i1.26380

[34] Loukina, A., Zechner, K., Chen, L., Leong, C. W., Ahmed, B., & Madnani, N. (2021). Artificial intelligence in language assessment. *Annual Review of Applied Linguistics*, 41, 111-126.

[35] Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). Intelligence unleashed: An argument for AI in education. Pearson.

[36] Luckin, R., Hutchinson, J., du Boulay, B., & Livingstone, I. (2021). Human intelligence, artificial intelligence and the allied forces Beshke of education systems fostering human-AI collaboration and co-evolution. *British Journal of Educational Technology*, 52(4), 1361-1373.

[37] Madnani, N., Loukina, A., Von Davier, A., Burstein, J., & Cahill, A. (2018, April). Building better open-domain conversational agents by using dialog acts. In Proceedings of the Workshop: Analyzing and Integrating Conversational Strategies in Dialogue Systems (pp. 43-48).

[38] Ma, W., Adesope, O. O., Nesbit, J. C., & Liu, Q. (2014). Intelligent tutoring systems and learning outcomes: A meta-analysis. *Journal of Educational Psychology*, 106(4), 901.

[39] Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers college record*, 108(6), 1017-1054.

[40] Nye, B. D., Graesser, A. C., & Hu, X. (2014). AutoTutor and family: A review of 17 years of natural language tutoring. *International Journal of Artificial Intelligence in Education*, 24(4), 427-469.

[41] Qian, L., Yang, Y., & Zhao, Y. (2021). Syntactic complexity revisited: sensitivity of China's AES-generated scores to syntactic measures, effects of discourse mode and topic. Read. Writ. 34(3), 681–704. https://doi.org/10.1007/s11145-020- 10087-5

[42] Radanliev, P., & De Roure, D. (2022). Advancing the cybersecurity of the healthcare system with self-optimizing and self-adaptative artificial intelligence (part 2). *Health and Technology*, 12(5), 923-929. https://doi.org/10.1007/s12553-022-00691-6

[43] Radanliev, P., & De Roure, D. (2022). New and emerging forms of data and technologies: Literature and bibliometric review. *Multimedia Tools and Applications*, 1-25. https://doi.org/10.1007/s11042-022-13451-5

[44] Roberts-Mahoney, H., Means, A. J., & Garrison, M. J. (2016). Netflixing human capital development: Personalized learning technology and the corporatization of K-12 education. *Journal of Education Policy*, 31(4), 405-420.

[45] Selwyn, N. (2019). Should robots replace teachers? AI and the future of education. John Wiley & Sons.

[46] UNSW Report. (2020). Artificial Intelligence for Inclusive Education: A Review of Theory, *Practice and Policy*.

[47] Woolf, B. P., Lane, H. C., Chaudhri, V. K., & Kolodner, J. L. (2013). AI grand challenges for education. *AI Magazine*, 34(4), 66-84.

[48] Zhang, M. (2021). Evaluating automated writing evaluation tools: A case study of ETS® Criterion®. *Language Testing in Asia*, 11(1), 1-21.

[49] Zitouni, K. S. (2022). New trends in EFL online learning and teaching through the lens of artificial intelligence. *Almuqadimah Hum. Soc. Stud. J.* 7(1), 1065–1080.

Appendix

Here is a draft Likert scale questionnaire that could be used to collect quantitative data from EFL teachers covering the key sections:

Section 1: Experience with AI Tools

1. I am familiar with AI technologies used in EFL learning.

2. I frequently integrate AI into my EFL lessons.

Section 2: Perceptions of AI's Pedagogical Role

3. AI helps personalize instruction to students' needs.

4. AI improves practice opportunities for language skills.

5. AI risks reducing meaningful interaction between students and teachers.

Section 3: Collaboration with AI Systems

6. AI plays a supportive role in my classroom without replacing human teachers.

7. AI helps improve efficiency by automating routine tasks.

8. Teachers should maintain control over content delivered via AI.

Section 4: Privacy, Bias and Fairness

9. I am concerned about student data privacy when AI systems are used.

10. AI systems may unintentionally discriminate against some groups of students.

11. Guidelines are needed regarding appropriate and fair use of AI for learning.

Section 5: Assessment and Evaluation

12. AI shows promise for standardized scoring of student work like essays.

13. AI evaluations of students align with expert human judgment.

14. More validation of AI assessment models is required.

Section 6: Support for Effective Integration

15. I received adequate training on integrating AI into my teaching.

16. Technical and instructional support is available when issues arise.

17. My role as a teacher enhances rather than replaces when partnering with AI.

Scale: 1 (Strongly Disagree) to 5 (Strongly Agree)

Open comment box at the end to collect additional viewpoints.