*Original Research Article*

Towards Responsible AI Ethical Utilization: Exploring Ethical Guidelines of AI-generated EFL Formative assessments

.

ABSTRACT

|  |
| --- |
| **Aims:** This exploratory mixed-methods study investigates Saudi EFL teachers’ perceptions of ethical guidelines related to the use of AI-generated formative assessments. It also examines the ethical challenges they face when integrating AI into EFL classroom practices.  **Study Design:** An explanatory sequential mixed-methods design was employed. The study began with quantitative data collection through a questionnaire administered to Saudi EFL teachers, followed by qualitative data collection via semi-structured interviews to elaborate on and contextualize the survey findings.  **Place and Duration of Study:** The research was conducted at the English Language Institute, King Abdulaziz University, between September and December 2024.  **Methodology:** The sample consisted of 39 Saudi EFL teachers (22 males and 17 females), aged between 20 and over 51, with varying educational backgrounds and teaching experience. Quantitative data were analyzed using SPSS, and qualitative interview data from three teachers (two males, one female) were thematically analyzed to identify recurring ethical concerns.  **Results:** Descriptive statistics showed gender-based differences in responses. A t-test (t = -18.15 < 2.042) indicated a significant difference in one section of the questionnaire, possibly due to differences in experience or perceptions. However, the second section reflected broadly aligned views between male and female participants.  **Conclusion:** Findings indicate that while teachers acknowledge the potential benefits of AI in enhancing formative assessment, they lack adequate ethical guidance and institutional support for its responsible use. Ethical concerns and practical constraints currently limit the effective adoption of AI tools in formative assessment practices. The study contributes to the field of AI in language education by emphasizing the need for clear, context-sensitive ethical frameworks and institutional policies to support EFL teachers in responsibly integrating AI into formative assessment. |

*Keywords: {Artificial Intelligence, Ethical consideration, formative assessments, Saudi English teachers }*

1. INTRODUCTION

Recently, the emergence of new technologies has significantly transformed the realm of education, particularly in English language teaching, with the current trend focusing on the integration of artificial intelligence (AI) into learning environments. AI is defined as the ability to closely mimic the human’s brain’s activities to achieve a certain goal using machines.

Research studies have proven the advantages and disadvantages of AI when implemented in English as a Foreign Language (EFL) classroom. It greatly enhances teachers’ efficacy, reduces their workload, and assists them in generating personalized materials and assessments to improve students’ different linguistic skills (Ali, 2023; Hwang et al., 2020; & Haung; 2022). Despite these advantages, major concerns have been associated with the implementation of AI in EFL classrooms. Researchers have found that English teachers are mostly worried about the ethical challenges that they face, including privacy invasion, inaccurate generated materials, unfairness, and lack of explainability (Chounta et al., 2022; Hartono et al., 2023; Simbolon, 2024).

The origins of AI ethics in education can be traced back to pioneers such as Aiken and Epstein, who published one of the earliest papers on ethical guidelines for AI in education (Latham, Annabel & Goltz, Sean, 2019). Writing in 2000, they predicted that by 2010, AI would become a dominant force in education, reshaping teachers' roles significantly. Although their perspective was optimistic, they maintained a cautionary tone regarding the ethical implications of AI use in educational settings. Aiken and Epstein (2000) warned that AI could negatively affect young learners’ ethics, aesthetics, physical well-being, emotions, intellect, and social skills. While their work predates contemporary generative AI tools like ChatGPT, their emphasis on the ethical risks of automation in education remains relevant. Their article laid the theoretical groundwork for ongoing debates, even though it primarily focused on students rather than educators. It also lacked discussion of institutional responsibility or practical guidance for real-world implementation—gaps that are increasingly critical today as AI becomes integrated into classroom practices. By revisiting this foundational work, this study builds on its ethical concerns while extending the conversation to include teachers' perspectives and institutional roles in the age of advanced AI.

Ethical issues related to the utilization of AI could affect the educational process negatively. It can lead to the improper use of AI tools or result in negative outcomes. The generated content can sometimes be inaccurate, potentially causing spreading misleading information to the students (Chan & Hu, 2023). Moreover, teachers may use students’ data and assessments to generate personalized content, which could be mishandled or used without consent, jeopardizing their privacy and security. Furthermore, teachers may struggle in explaining the decisions made by AI, leading to a lack of trust and acceptance of the data generated. This could also hinder their ability to justify their choices within the classroom which reduces their confidence in their practices (Wang, Bian, & Chen, 2024).

Because of the negative consequences of the ethical concerns that are associated with the utilization of AI, initiatives have been undertaken to address them. They aim to design clear guidelines for the appropriate use of AI. For instance, the Ministry of Science and Technology in China released “Ethical Rules for New-generation Artificial Intelligence” in 2021 as an example. It stresses on identifying ethical principles underlying AI technology to reduce its danger towards humanity (Haung, 2022). Additionally, UNESCO (2022) states that it is important for AI systems to respect and advance human rights, freedoms, and dignity at all costs.

In Saudi Arabia, the Saudi Data and AI Authority (SDAIA) has issued a guide to the accurate use of AI in alliance with Saudi Arabia’s cultural and religious values. Based on the published ethical guidelines relevant to the educational settings, there are concerns that are prominently featured in most guidelines and align with SDAIA’s standards.

One of the most prominent ethical challenges is privacy. Privacy concerns arise when AI tools require access to sensitive student data to provide personalized feedback in formative assessments. While consent mechanisms exist, users often approve data access without full awareness of the implications. In EFL settings, student writing, progress, and language errors are regularly submitted to AI platforms. Without strict institutional data policies, this poses a risk of privacy invasion, particularly as data storage and use may fall outside the teacher’s control (Vincent-Lancrin & Van der Vlies, 2020). Teachers need clear guidelines on how student data is collected, stored, and protected when using AI for assessment.

Additionally, explainability—the ability to understand and justify AI decisions—is critical in EFL formative assessment. Since AI tools often function as “black boxes,” teachers may find it difficult to comprehend how feedback is generated (Goebel et al., 2018). In language education, where feedback must align with pedagogical goals and linguistic accuracy, this lack of transparency undermines teacher trust and confidence. When teachers understand AI decisions and can link them to learning theories, they are more likely to responsibly integrate AI outputs into classroom practice (Theodorou et al., 2017).

Furthermore, accountability refers to identifying who is responsible for AI-generated errors. In EFL assessment, this becomes problematic when AI provides incorrect feedback or biased scoring. Some scholars argue that developers are responsible (Santoni de Sio & Mecacci, 2021), while others place accountability on teachers and institutions for failing to validate AI outputs (Vincent-Lancrin & Van der Vlies, 2020). Ultimately, educators cannot delegate responsibility to AI; they must verify the accuracy of the feedback to protect learner outcomes and maintain professional integrity.

Bias in AI tools can result in discriminatory outcomes, especially in language evaluation. For example, studies show AI may recommend different courses based on gender or score essays lower when written in non-standard dialects (Khan, 2023). These biases can demotivate learners and unfairly influence teacher judgments in formative assessment. Moreover, unequal access to AI tools—due to limited digital infrastructure or varying student digital literacy—raises concerns about equity in EFL classrooms (Nye, 2015). Teachers must be aware of these risks and adopt strategies to mitigate their impact.

Accuracy is essential for AI to support formative assessment effectively. AI tools often generate plausible but incorrect feedback due to the limitations of their training data (Sallam, 2023). In EFL contexts, where feedback guides language development, misleading content can harm student progress. Research shows that teachers and students remain skeptical of AI accuracy, particularly when fabricated references or vague language explanations are detected (Aguilera-Hermida, 2024; Chan & Hu, 2023). Teachers must critically evaluate AI-generated responses before incorporating them into instruction.

Moreover, reliability involves the consistency and objectivity of AI-generated outputs, while trustworthiness reflects user confidence in the tool’s ability to perform effectively. In EFL placement and writing assessments, AI tools like ChatGPT have demonstrated moderate to low scoring reliability (Kim et al., 2024), failing to account for topic specificity or source integration. These limitations highlight the ethical obligation to scrutinize AI-generated assessments before acting on them. Trustworthy AI depends not only on consistent performance but also on explainability, privacy safeguards, and institutional oversight (OECD, 2021; Nguyen et al., 2022).

Saudi Arabia’s 2030 vision emphasizes the importance of digital development especially in education. Therefore, various Saudi teachers implement different AI tools, such as ChatGPT and Poe in their practices. Xu (2022) explains that for educators to utilize AI effectively, they are obliged to be aware of the AI-based tools’ pedagogical contributions. Additionally, for teachers to select the appropriate AI tools, they need to be knowledgeable about them. Furthermore, ethics plays a crucial part in the process of effective implementation. The constant implementation and focus on the advantages have led to the negligence of the severe drawbacks that these tools may cause. Al-Zahrani (2017) illustrates that English teachers lack major competencies in relation to the ethical and accurate use of these tools. As a result, unwanted outcomes may emerge, leading to harmful effects for both teachers and students.

Although some empirical studies do not explicitly focus on ethics, they reveal that teachers are increasingly aware of the ethical implications of AI in classrooms. For instance, Hartono et al. (2023) report that while teachers generally express positive attitudes toward AI, they also cite insufficient training and support as barriers to effective implementation. This indirectly reflects a key ethical concern—explainability—as teachers require not only technical skills but also the capacity to interpret and justify AI-generated outputs within pedagogical contexts. In the Saudi and broader GCC context, however, there is a notable lack of empirical research examining teachers' perceptions of AI ethics in education. One study by Shamsuddinova, Heryani, and Naval (2024), conducted across several GCC countries, explored 11 educators’ views on AI implementation. While participants expressed optimism, they also underscored significant concerns, such as limited understanding of AI tools and the absence of localized ethical guidelines. The small sample size reflects the infancy of this research area in the region, but it also signals a critical gap: why have educational institutions in the GCC not developed clear, context-specific AI ethical frameworks despite rapid digitalization? This gap may be attributed to multiple factors, such as the top-down nature of policymaking in many GCC countries, limited interdisciplinary collaboration between educators and AI developers, or a focus on infrastructure over pedagogical ethics. The absence of institutional guidelines not only limits the responsible adoption of AI but also places the ethical burden disproportionately on individual teachers, many of whom lack the training or authority to navigate these complex issues independently. Therefore, there is an urgent need for GCC-specific ethical frameworks that account for regional values, educational goals, and institutional capacities.

The Technological, Pedagogical, and Content Knowledge (TPACK) framework comprises three intersecting domains: technological content knowledge (TCK), technological pedagogical knowledge (TPK), and technological pedagogical content knowledge (TPACK). Due to its comprehensive and adaptable structure, TPACK has been widely used in studies examining the integration of various technologies in educational settings (Mishra et al., 2010). More recently, researchers have attempted to link AI ethics to the TPACK framework to better address the ethical complexities emerging from AI integration in classrooms. Celik (2023) introduced the “Intelligent TPACK” (iTPACK) model, an extension of the original framework that incorporates ethical considerations as an additional component. Unlike treating ethics as an embedded feature of TPK or TCK, Celik positions ethics as a distinct domain, focusing primarily on transparency, accountability, fairness, and inclusiveness. However, the model overlooks other crucial ethical issues such as privacy, explainability, reliability, and trustworthiness, which are particularly significant when using AI tools in educational assessments. Furthermore, although the framework emphasizes ethics conceptually, this emphasis is not strongly reflected in the empirical instrument. For instance, while components like Intelligent-TK, Intelligent-PK, and Intelligent-CK each contain over ten items in the questionnaire, the ethics domain includes only four, suggesting a relatively underdeveloped focus on ethical literacy. In contrast, Gómez-Trigueros (2023) offers an alternative ethics-focused adaptation of the TPACK model, proposing a framework explicitly designed to evaluate teachers' ethical knowledge in technology use. Her version includes variables such as Professional Ethical Knowledge (PEK), Technological Ethical Knowledge (TEK), Pedagogical Technological Ethical Knowledge (PTEK), and Ethical Disciplinary Technology Knowledge (EKDT)**,** which together provide a more granular view of ethics in educational technology.

This study adopts the TPACK framework to explore Saudi EFL teachers’ perceptions of ethical guidelines and the challenges they face in using AI-generated formative assessments. TPACK's flexibility allows for adaptation to this specific context, and its holistic structure—encompassing pedagogy, content, and technology—makes it well-suited for evaluating the multifaceted ethical considerations AI introduces. Drawing on prior adaptations by Celik and Gómez-Trigueros, the current study modifies the TPACK framework to emphasize the ethical dimensions of AI use, aligning with the research objective of identifying practical and ethical challenges in Saudi EFL classrooms.

Assessments usually care for monitoring the learners’ progress throughout the whole learning process. There are two types of assessments: formative and summative. Black and William (1998) define formative assessments as engaging in-class assessments that inform teachers’ instructional and learning decisions on students’ communication. As a result, the main goal is to enhance the learners’ learning experiences. Observations, performance tasks, and portfolios are all examples of formative assessments. Mostly, these assessments are not graded. In the context that this research investigates, which is the ELI institution in the King Abdulaziz University, there is a defined sets of procedures that assessments follow, therefore, English teachers have limited control over the assessment creating process. Mansory (2019) explains in his dissertation that each semester students undergo a placement test, a continuous portfolio, two writing exams, two speaking exams, a mid-module exam, and a final exam. Most exams are administrated by the ELI testing Committee, thus teachers have no role in designing these exams. However, English teachers are in charge of in-class ungraded formative assessments that monitor the learners’ progress throughout the course. Therefore, to achieve the aim of this research, AI-generated formative assessments would be the highlight. Although there is limited research specifically addressing the direct impact of AI tools on English assessments, Keerthiwansha (2018) explains that a speaking assessment could take approximately four hours, negatively impacting both teachers and students. However, with the integration of AI, teachers can use student data to create specific assessments based on individual needs and utilize AI tools to administer the tests. Moreover, these tools allow teachers to receive not only grades but also detailed feedback on errors, enabling a targeted focus in future assessments. Additionally, using AI tools in designing English assessments improves students’ classroom achievement. However, in order to ensure ethical implementation, teachers need to be aware of the ethical concerns that may arise while generating EFL formative assessments, such as inaccurate and bias-generated assessments, students’ data invasion, and not being able to explain the generated assessments. To achieve that, Tapalova and Zhiyenbayeva (2022) suggest the necessity of training teachers to adapt to the new technological advancements particularly in relation to understanding AI tools thoroughly and avoiding excessive dependence. Similarly, Cassidy (2023) emphasizes the importance of educating teachers regarding AI ethics to ensure an effective and responsible utilization of it.

Due to the growth of utilizing AI in educational contexts, especially in English language classrooms, there is a need to draw educators’ attention to the ethical guidelines associated with this implementation. The reason for that is to ensure the effective implementation of AI. As discussed in previous studies in different contexts, teachers perceive AI positively because of its advantages. However, most previous studies have not examined the ethics of AI in depth. Instead, some ethical concerns have been addressed while investigating the teachers’ perception of AI tools. Additionally, the focus of most perception studies was mainly on the advantages and disadvantages of AI tools. Therefore, this research aims to address the following research questions: 1) What is Saudi English language teachers’ perception toward utilizing AI ethical guidelines for EFL formative assessments? 2) What are the challenges that EFL Saudi teachers encounter in utilizing AI ethical guidelines for formative assessments?

2. methodology

This study adopted an explanatory sequential mixed-methods design (Creswell & Creswell, 2018), which integrates quantitative and qualitative phases to provide both breadth and depth. The research aimed to examine Saudi EFL teachers’ perceptions of ethical guidelines in AI-generated formative assessments, as well as the ethical challenges they encounter when implementing these tools in their classrooms. The design began with the quantitative collection and analysis of survey data, followed by qualitative interviews to explain and elaborate on the statistical findings.

**2.1 Research Design**

The explanatory sequential mixed-methods approach was selected for its ability to produce generalizable patterns through the quantitative phase while capturing contextual richness in the qualitative phase (Teddlie & Tashakkori, 2009). This design was especially relevant given the emerging nature of AI ethics in Saudi educational contexts, where both measurable trends and deeper insight are essential.

**2.2 Participants**

Participants in the quantitative phase included 39 Saudi EFL teachers from the English Language Institute (ELI) at King Abdulaziz University in Jeddah. Inclusion criteria required participants to be Saudi nationals currently teaching English at the university level. A convenience sampling strategy was employed based on participant availability and willingness to contribute.

For the qualitative phase, five participants were selected from the survey respondents who volunteered to be interviewed. These individuals were chosen to ensure representation across teaching experience levels and gender. Their willingness was indicated via a consent section embedded in the questionnaire.

**2.3 Quantitative Data Collection and Analysis**

A 29-item questionnaire adapted from Gómez-Trigueros (2023)—originally grounded in the TPACK framework—served as the primary data collection tool. Items were revised to reflect the ethical dimensions of AI use in EFL formative assessment, ensuring cultural and contextual relevance. The scale employed a 5-point Likert format (1 = Strongly Disagree, 5 = Strongly Agree), with no neutral option to reduce ambiguous responses, in line with Johns (2010).

The questionnaire consisted of three sections:

1. Demographic information.
2. Ethical perceptions of AI-generated formative assessment.
3. Analysis of TPACK in relation to AI ethics
4. Consent and willingness to participate in follow-up interviews.

Data was collected online using Google Forms, allowing broad and efficient access. Ethical approval was secured from the ELI's research committee, and informed digital consent was obtained. The instrument’s content validity was established through expert reviews by seven PhD scholars specializing in AI ethics and EFL pedagogy. Cronbach’s Alpha confirmed high internal reliability (Table 1).

**Table 1. Cronbach’s Alpha Coefficients for Questionnaire Constructs**

| **Constructs** | **Corresponding Items** | **Cronbach’s Alpha** |
| --- | --- | --- |
| Knowledge about AI Ethics | 1–18 | 0.93 |
| TPACK-Based Ethical Knowledge | 19–29 | > 0.67 |

Quantitative data was analyzed using SPSS Version 24. Descriptive statistics including frequencies, means, and standard deviations were used to summarize teacher perceptions. These statistics helped identify the general trends and potential ethical concerns perceived by teachers, which were explored further in the interviews.

**2.4 Qualitative Data Collection and Analysis**

To gain deeper insight, semi-structured interviews were conducted with five participants. The interview protocol was informed by the literature (Rubin & Rubin, 2012; Kvale, 2015) and findings from the quantitative phase. Questions were organized around four themes:

1. Teachers’ experience with AI.
2. Ethical challenges encountered.
3. Influence of cultural and religious factors.
4. Institutional support mechanisms.

Interviews were conducted via Zoom, lasting 30–35 minutes. Participants received the guide beforehand. All interviews were recorded with oral consent, transcribed verbatim, and analyzed using inductive thematic analysis, following Braun and Clarke’s (2006) six-step method. NVivo software facilitated coding and theme development.

Member checking was employed to verify data accuracy, and triangulation between quantitative and qualitative findings enhanced the trustworthiness of the results. Ethical safeguards included the use of pseudonyms, anonymized data storage, and documented informed consent at all stages**.**

3. results and discussion

This section presents both the quantitative findings and qualitative insights on Saudi EFL teachers’ perceptions of AI ethics in formative assessment. A total of 39 teachers participated in the study—57.89% male and 42.10% female—with varied age groups, academic qualifications, and teaching experience. Most held a master's degree and had 11 to 30 years of teaching experience.

**3.1. AI Use in Formative Assessment**

Only 27.95% of participants reported using AI to generate formative assessments, while 72.05% did not (Table 1). Most of the AI users applied the technology rarely rather than frequently or consistently (Table 2).

**Table 2. AI Usage Among Participants**

| **Response** | **Male** | **Female** | **Percentage** |
| --- | --- | --- | --- |
| Yes, I integrate AI | 8 | 2 | 27.95% |
| No, I do not integrate AI | 14 | 13 | 72.05% |
| **Total** | 22 | 15 | 100% |

**Table 3. Frequency of AI Use for EFL Formative Assessment**

| **Frequency** | **Male** | **Female** |
| --- | --- | --- |
| Frequently | 3 | 1 |
| Rarely | 5 | 1 |
| Never | 0 | 0 |

This limited integration supports findings by Hartono et al. (2023), who noted that despite teachers’ optimism, lack of training hinders effective AI adoption.

**3.2. Formal Training in AI Ethics**

Only 34.21% of participants had received formal training on AI ethical guidelines, with the remainder relying on informal sources such as workshops, webinars, social media, and self-study (Table 3). This highlights a significant gap in institutional preparation.

**Table 4. Formal Training on AI Ethics**

| **Response** | **Male** | **Female** | **Percentage** |
| --- | --- | --- | --- |
| Yes | 6 | 6 | 34.21% |
| No | 16 | 9 | 65.79% |
| **Total** | 22 | 15 | 100% |

This aligns with regional findings by Shamsuddinova et al. (2024), who emphasized the absence of structured, context-specific ethical guidelines in the GCC, leading to inconsistent practices and varying levels of teacher preparedness.

**3.3. Perceptions of General AI Ethical Guidelines**

Table 5 presents mean and standard deviation (SD) values across several items. The Likert scale was coded such that 1 = Strongly agree and 4 = Strongly disagree; thus, lower mean values indicate stronger agreement.

**Table 5. Sample Items on General AI Ethics Knowledge**

| **Item** | **X (Male)** | **SD (Male)** | **X (Female)** | **SD (Female)** |
| --- | --- | --- | --- | --- |
| It is crucial English teachers understand AI ethical guidelines. | 1.52 | 0.89 | 1.33 | 0.48 |
| I know bias can exist in AI algorithms. | 1.65 | 0.93 | 2.20 | 0.86 |
| I know obtaining consent before inserting data into AI is important. | 1.73 | 0.91 | 2.00 | 0.75 |

Despite both groups showing positive perceptions overall, the gender-based differences were significant. Female teachers showed greater internal consistency (lower SDs), while male teachers had more diverse perceptions. These variations may stem from differing exposure to technology or institutional training opportunities. This gender gap may reflect differences in access to digital resources or professional development opportunities, echoing Karimov et al. (2024), who found that male educators often have greater digital confidence.

**3.4. TPACK-Based Ethical Knowledge**

The second part of the questionnaire examined perceptions using the TPACK framework. Both male and female teachers expressed moderate agreement with statements concerning technical, pedagogical, and content-related ethical considerations. Differences were most visible in questions about awareness of bias and staying current with AI developments (Table 6).

**Table 6. TPACK Ethical Perception (Selected Items)**

| **Item** | **X (Male)** | **SD (Male)** | **X (Female)** | **SD (Female)** |
| --- | --- | --- | --- | --- |
| Misusing AI jeopardizes data security. | 1.26 | 0.44 | 1.86 | 0.51 |
| I keep up with AI developments in formative assessments. | 2.13 | 1.01 | 2.26 | 0.79 |
| Teachers should guide students in understanding AI's role in assessment. | 1.65 | 0.93 | 1.66 | 0.61 |

These findings suggest that while ethical intent is high, practical knowledge and ongoing learning about AI remain underdeveloped among Saudi EFL teachers.

**3.5. Inferential Statistics and Gender Differences**

To assess whether gender differences in responses were statistically significant, t-tests were conducted on both sections of the questionnaire.

* General AI Ethics: A significant difference was found between male (X = 1.89, SD = 1.11) and female (X = 2.01, SD = 1.24) teachers, t(30) = -18.15, *P* < .001. This indicates that male participants reported a more favorable perception of AI ethics.
* TPACK Ethics: No significant gender difference was observed in TPACK-related responses, t(18) = -1.05, *P* = .308. This suggests that both genders were similarly aligned in their perceptions of integrating AI ethically in instructional practice.

**3.6 Open-ended question analysis**

The first section of the questionnaire concluded with an open-ended question asking participants who should be held accountable for inaccurate AI-generated EFL formative assessments. Responses varied. Twenty out of 37 participants believed that teachers should be held accountable. One participant explained, *“I believe this only can happen if the information entered into AI is not accurate nor clear.”* Another stated, *“If the input is accurate, the data will be accurate... we can say ‘whoever put the input.’”* These responses suggest that many participants view teachers as primarily responsible, given their role in crafting and entering prompts.

Other participants, however, attributed responsibility to different stakeholders, including school administration, AI systems themselves, and AI developers. A few participants did not provide a response, simply writing “N/A” or “I don’t know.” This range of views reflects broader global debates about AI accountability in education, as highlighted by Santoni de Sio and Mecacci (2021), who emphasize the complexity of distributing ethical responsibility among users, institutions, and developers.

**3.7 Findings**

The findings of this study provide critical insights into Saudi English teachers’ perceptions of AI ethics and the challenges they face in implementing ethical guidelines during AI-assisted formative assessment.

Overall, teachers displayed positive perceptions of AI-generated formative assessments, with male teachers showing significantly stronger agreement than females. This gender-based divergence may reflect differing digital confidence levels or access to AI-related training, consistent with Karimov et al. (2024). Interestingly, despite differences in individual item responses, both genders agreed on the importance of ethical awareness, especially in regard to accuracy, fairness, and privacy.

The data also showed a mixed understanding of AI ethics. While many teachers mentioned concepts like transparency and data protection, several equated AI ethics with research ethics. Although overlapping in values, the two domains differ in application and scope. This confusion suggests a need for clearer training that delineates AI ethics as a standalone professional competency.

Participants reported varying degrees of access to formal training on AI ethics, even within the same institution. This discrepancy points to inconsistencies in the internal dissemination of professional development and supports Hartono et al.'s (2023) argument for centralized and equitable access to AI ethics training.

The lack of uniformity not only affects ethical knowledge but also contributes to inequality in implementation, with some teachers improvising while others operate with institutional support. This may widen gaps in assessment quality and fairness across classrooms.

The qualitative analysis revealed three core ethical concerns: accuracy, dataprivacy, and bias. Teachers were cautious about AI’s reliability, especially in grading and content generation, echoing Harrer’s (2023) concerns about AI’s pedagogical limitations. Similarly, fears regarding data privacy, particularly with tools like ChatGPT, were consistent with findings from Vincent-Lancrin & Van der Vlies (2020).

Participants also highlighted concerns over accountability. In open-ended responses, they identified various responsible entities—teachers, institutions, and AI developers—mirroring global debates (Santoni de Sio & Mecacci, 2021). This ambiguity emphasizes the need for clear accountability frameworks within institutional AI policies.

These findings underscore the need for:

* **Comprehensive AI ethics training** programs targeting digital literacy, privacy, and bias.
* **Institutional guidelines** aligning local policies with global AI ethics frameworks
* **Localized development** of AI tools that are culturally sensitive and pedagogically aligned.
* **Transparent AI systems** to support teacher understanding and student trust.

**3.8 Practical Implications**

Given these findings, Saudi institutions should take immediate steps to:

* **Develop formal AI ethics training** for EFL educators, with emphasis on accuracy, transparency, consent, and cultural context.
* **Implement standardized, localized ethical guidelines** aligned with global frameworks but adapted to GCC-specific educational, cultural, and religious contexts.
* **Promote transparent and explainable AI tools**, aiding both teacher understanding and student trust.
* **Encourage institutional accountability**, clearly designating roles and responsibilities in AI use across educational layers.

By addressing these systemic gaps, Saudi educational institutions can empower teachers to use AI ethically, fairly, and effectively ensuring formative assessments remain valid, equitable, and trustworthy.

4. Conclusion

This study explored the challenges that Saudi EFL (English as a Foreign Language) teachers face in applying ethical guidelines when using artificial intelligence (AI) for formative assessment. Guided by the central research question—*What are the challenges that EFL Saudi teachers encounter in utilizing AI ethical guidelines for formative assessments?*—the research uncovered both practical and conceptual barriers that hinder ethical AI implementation in Saudi classrooms.

Using a mixed-methods approach, the study drew on survey and interview data to provide a holistic understanding of teachers’ experiences. The findings revealed a general openness toward integrating AI into formative assessment; however, this enthusiasm was constrained by a lack of formal training, limited awareness of ethical principles, and insufficient institutional support. Teachers expressed particular concern over issues such as data privacy, algorithmic bias, cultural alignment, and accountability.

This research makes several key contributions. First, it highlights the ethical literacy gap among Saudi EFL teachers in relation to AI use. Second, it underscores the absence of localized ethical frameworks, revealing a disconnect between rapid technological adoption and institutional preparedness. Third, it provides empirical evidence that supports the need for context-sensitive training and policy reform, aligning directly with the educational goals outlined in Saudi Vision 2030, which emphasizes innovation, digital transformation, and human capital development.

* 1. **Implications for Practice**:

To translate these findings into practice, several actionable recommendations emerge:

* **Professional Development**: Ethical AI literacy should be embedded in teacher training programs. In-service modules should cover core principles such as data privacy, transparency, fairness, and accountability in AI use for assessment.
* **Localized Ethical Guidelines**: Institutions should collaborate with educators, technologists, and ethicists to create context-specific AI policies that are sensitive to Saudi cultural and pedagogical values.
* **AI Tool Vetting**: Ministries and educational institutions should provide access to approved AI tools that meet ethical and educational standards, ensuring alignment with curriculum goals and data governance policies.
* **Accountability Structures**: Clear frameworks are needed to assign responsibility when ethical breaches occur, whether due to teacher misuse, system error, or administrative oversight.

These steps are essential to ensuring that AI enhances rather than undermines the equity, reliability, and trustworthiness of formative assessment.

* 1. **Directions for Future Research**:

Future research could investigate the perspectives of students regarding AI-driven formative assessment, offering a more balanced understanding of its ethical impact. Longitudinal studies could track how teacher attitudes and ethical practices evolve over time with sustained training. Comparative studies between public and private institutions may also yield further insights into contextual differences in AI implementation.

The integration of AI into formative assessment presents both promise and peril. This study highlights the critical importance of ethical literacy and institutional support in ensuring that technological adoption enhances rather than undermines educational equity, trust, and effectiveness.

* 1. **Limitations of the Study**

While the mixed-methods approach enriched the findings, several limitations must be acknowledged. The study was geographically limited to select institutions within Saudi Arabia, potentially limiting the generalizability of its results. Furthermore, the reliance on self-reported data in surveys and interviews may introduce bias or reflect social desirability.

These limitations may have influenced the breadth and depth of the data, particularly in relation to participants’ openness in discussing institutional shortcomings or ethical dilemmas. Future studies could employ observational data or longitudinal tracking to overcome these constraints

Consent

Written informed consent was obtained from the participants for publication of this research. A copy of the written consent is available for review by the Editorial office/Chief Editor/Editorial Board members of this journal.

Disclaimer (Artificial intelligence)

Option 1:

Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc.) and text-to-image generators have been used during the writing or editing of this manuscript.

Option 2:

Author(s) hereby declare that generative AI technologies such as Large Language Models, etc. have been used during the writing or editing of manuscripts. This explanation will include the name, version, model, and source of the generative AI technology and as well as all input prompts provided to the generative AI technology

Details of the AI usage are given below:

1.AI is used to generate synonyms and paraphrase some sentences.

2.AI is used to edit grammatical errors and sentence strutures.

References

1. Ali, A. M. (2023). An Intervention Study on the Use of Artificial Intelligence in the ESL Classroom: English teacher perspectives on the Effectiveness of ChatGPT for Personalized Language Learning. [Master Thesis, Malmo University] Digitala Vetenskapliga Arkivet. <https://www.diva-portal.org/smash/get/diva2:1774035/FULLTEXT02>
2. Huang, L. (2022). An empirical study of integrating information technology in english teaching in artificial intelligence era. *Scientific Programming*, *2022*(1), 6775097. <https://onlinelibrary.wiley.com/doi/pdf/10.1155/2022/6775097>
3. Hwang, G. J., Xie, H., Wah, B. W., & Gašević, D. (2020). Vision, challenges, roles and research issues of Artificial Intelligence in Education. *Computers and Education: Artificial Intelligence*, *1*, 100001. <https://www.sciencedirect.com/science/article/pii/S2666920X20300011>
4. Chounta, I. A., Bardone, E., Raudsep, A., & Pedaste, M. (2022). Exploring teachers’ perceptions of artificial intelligence as a tool to support their practice in Estonian K-12 education. *International Journal of Artificial Intelligence in Education, 32*(3), 725-755.<https://www.researchgate.net/profile/Irene-Angelica-Chounta/publication/352066912_Exploring_Teachers%27_Perceptions_of_Artificial_Intelligence_as_a_Tool_to_Support_their_Practice_in_Estonian_K-12_Education/links/60b884a6a6fdcc22eacf5829/Exploring-Teachers-Perceptions-of-Artificial-Intelligence-as-a-Tool-to-Support-their-Practice-in-Estonian-K-12-Education.pdf>
5. Hartono, W. J., Nurfitri, N., Ridwan, R., Kase, E. B., Lake, F., & Zebua, R. S. Y. (2023). Artificial Intelligence (AI) Solutions In English Language Teaching: Teachers-Students Perceptions And Experiences. *Journal on Education*, *6*(1), 1452-1461. <http://jonedu.org/index.php/joe/article/download/3101/2642>
6. Simbolon, T. N. (2024). Bridging The Gap: Language Teacher Understanding and Responsible AI Implementation in Schools. A Case Study of English Language Educators in Kubu Raya. *Journal on Education*, *6*(3), 16057-16068.‏ <https://www.jonedu.org/index.php/joe/article/download/5488/4394>
7. Aiken, R. M., & Epstein, R. G. (2000). Ethical guidelines for AI in education: Starting a conversation. *International Journal of Artificial Intelligence in Education*, *11*(2), 163-176.
8. Latham, A., & Goltz, S. (2019). A Survey of the General Public’s Views on the Ethics of Using AI in Education. In *Artificial Intelligence in Education: 20th International Conference, AIED 2019, Chicago, IL, USA, June 25-29, 2019, Proceedings, Part I 20* (pp. 194-206). Springer International Publishing. <https://e-space.mmu.ac.uk/622786/3/Ethics%20in%20AIEDv5%20FINAL.pdf>
9. Chan, C. K. Y., & Hu, W. (2023). Students’ voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, *20*(1), 43.
10. Wang, D., Bian, C., & Chen, G. (2024). Using explainable AI to unravel classroom dialogue analysis: Effects of explanations on teachers' trust, technology acceptance and cognitive load. *British Journal of Educational Technology*.‏
11. SDAIA. (2023). *AI Ethics Principles*. SDAIA
12. Stahl, B. C., & Wright, D. (2018). Ethics and privacy in AI and big data: Implementing responsible research and innovation. *IEEE Security & Privacy, 16*(3), 26-33. <https://dora.dmu.ac.uk/bitstream/handle/2086/15328/SP_SPSI-2017-09-0274.R1_Stahl.docx?sequence=1&isAllowed=y>
13. UNESCO. (2022). Recommendation on the Ethics of Artificial Intelligence. <https://unesdoc.unesco.org/ark:/48223/pf0000381137>
14. Clarke, R. (1999). Internet privacy concerns confirm the case for intervention. *Communications of the ACM*, *42*(2), 60-67.
15. Vincent-Lancrin, S. and R. van der Vlies (2020), "Trustworthy artificial intelligence (AI) in education: Promises and challenges", *OECD Education Working Papers,* No. 218, OECD Publishing, Paris, <https://doi.org/10.1787/a6c90fa9-en>.
16. de Saint Laurent, C. (2018). In defence of machine learning: Debunking the myths of artificial intelligence. *Europe's journal of psychology*, *14*(4), 734. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6266534/>
17. European Commission, High-Level Expert Group on AI. (2019). *Ethics guidelines for trustworthy AI*. Brussels.
18. Goebel, R., Chander, A., Holzinger, K., Lecue, F., Akata, Z., Stumpf, S., ... & Holzinger, A. (2018). Explainable AI: the new 42?. In *International cross-domain conference for machine learning and knowledge extraction* (pp. 295-303). Springer, Cham.‏ <https://inria.hal.science/hal-01934928/file/cdMake18-XAI-Lecue-et-al.pdf>
19. Li, C., Xing, W., & Leite, W. (2024). Using fair AI to predict students’ math learning outcomes in an online platform. *Interactive Learning Environments*, *32*(3), 1117-1136. <https://www.tandfonline.com/doi/pdf/10.1080/10494820.2022.2115076?casa_token=xA1tygjeHMYAAAAA:JtMd1JkHoN5L6QKoxY3Zsvl3iZQBbTm-Z9ZL3Nc7J2N2Ez2bFv7s59Pyeu3KIHzjBAKHSkXzr78>
20. Memarian, B., & Doleck, T. (2023). Fairness, Accountability, Transparency, and Ethics (FATE) in Artificial Intelligence (AI), and higher education: A systematic review. *Computers and Education: Artificial Intelligence*, 100152.‏ <https://www.sciencedirect.com/science/article/pii/S2666920X23000310>
21. Santoni de Sio, F., & Mecacci, G. (2021). Four responsibility gaps with artificial intelligence: Why they matter and how to address them. *Philosophy & Technology, 34*(4), 1057-1084. <https://link.springer.com/content/pdf/10.1007/s13347-021-00450-x.pdf>
22. Shin, D., Rasul, A., & Fotiadis, A. (2022). Why am I seeing this? Deconstructing algorithm literacy through the lens of users. *Internet Research*, *32*(4), 1214-1234.‏ <https://www.emerald.com/insight/content/doi/10.1108/INTR-02-2021-0087/full/pdf?casa_token=UjzMv3CIWDwAAAAA:brozX99k-fRVdHAPMjgvOXU5Ab3il4RByqjG7pOjdtpJRfhXXz5HzJHIXpa5xs4TRBzyYgpUBUFYa8X1DjxyVpP3LVBttfP7BpWfJsl1eIp7EMKRhw>
23. Theodorou, A., Wortham, R. H., & Bryson, J. J. (2017). Designing and implementing transparency for real time inspection of autonomous robots. Connection Science, 29(3), 230–241. <https://doi.org/10.1080/09540091.2017.1310182>
24. Aguilera-Hermida, A. P. (2024, January). AI Chatbots in Education: The Importance of Accuracy. In *International Forum of Teaching & Studies, 20*(1).
25. Chan, C. K. Y., & Hu, W. (2023). Students’ voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, *20*(1), 43.
26. Fraunhofer IAIS.: Vertrauenswürdiger Einsatz von Künstlicher Intelligenz. (2019). <https://www.ki.nrw/wp-content/uploads/2020/03/Whitepaper_KI-Zertifizierung.pdf>
27. Harrer, S. (2023). Attention is not all you need: the complicated case of ethically using large language models in healthcare and medicine. *EBioMedicine*, *90*.
28. Khan, S. (2023). The Ethical Imperative: Addressing Bias and Discrimination in AI-Driven Education. *Social Sciences Spectrum*, *2*(1), 89-96. <https://sss.org.pk/index.php/sss/article/download/23/22>
29. Kim, H., Baghestani, S., Yin, S., Karatay, Y., Kurt, S., Beck, J., & Karatay, L. (2024). ChatGPT for Writing Evaluation: Examining the Accuracy and Reliability of AI-Generated Scores Compared to Human Raters. *Exploring artificial intelligence in applied linguistics*, 73-95. <https://www.iastatedigitalpress.com/plugins/books/154/format/1230/download/>
30. MI Garage: Ethics Framework. (2020). <https://www.digicatapult.org.uk/expertise/case-studies/study/machine-intelligence-garage/>
31. Nguyen, A., Ngo, H. N., Hong, Y., Dang, B., & Nguyen, B. P. T. (2023). Ethical principles for artificial intelligence in education. *Education and Information Technologies*, *28*(4), 4221-4241. <https://link.springer.com/content/pdf/10.1007/s10639-022-11316-w.pdf>
32. Nye, B.D. (2015). Intelligent Tutoring Systems by and for the Developing World: a review of trends and approaches for Educational Technology in a Global Context. *International Journal of Artificial Intelligence in Education, 25*, (2), 177-203.
33. OECD: Recommendation of the Council on Artifcial Intelligence. (2021). <https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449>
34. Rahayu, S. (2023). The Impact of Artificial Intelligence on Education: Opportunities and Challenges. *Jurnal Educatio FKIP UNMA*, *9*(4), 2132-2140.
35. Sallam, M. (2023, March). ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. In *Healthcare* (Vol. 11, No. 6, p. 887). MDPI.
36. Alzahrani, R. B. (2024). An Overview of AI Data Protection in the Context of Saudi Arabia. *International Journal for Scientific Research*, *3*, 1-20.<https://vsrp.co.uk/wp-content/uploads/9-IJSR-Vol.-3-No.-3-Mar-2024-Paper8-Ms.-Reema.pdf>
37. Hartono, W. J., Nurfitri, N., Ridwan, R., Kase, E. B., Lake, F., & Zebua, R. S. Y. (2023). Artificial Intelligence (AI) Solutions In English Language Teaching: Teachers-Students Perceptions And Experiences. *Journal on Education*, *6*(1), 1452-1461.<http://jonedu.org/index.php/joe/article/download/3101/2642>
38. Xu, L. (2020, December). The dilemma and countermeasures of AI in educational application. In *Proceedings of the 2020 4th international conference on computer science and artificial intelligence* (pp. 289-294). <https://dl.acm.org/doi/pdf/10.1145/3445815.3445863?casa_token=4aBXVJdWo1YAAAAA:aT6kcnAZXF69vlxpmnxfZCKYcGnUFsinD5NnS_KVB0XpE184e-DLG0Dt7f7PNvSfhR5w3_7eoi8>
39. Celik, I. (2023). Towards Intelligent-TPACK: An empirical study on teachers’ professional knowledge to ethically integrate artificial intelligence (AI)-based tools into education. *Computers in Human Behavior*, *138*, 107468.‏ <https://www.sciencedirect.com/science/article/pii/S0747563222002886>
40. Gomez-Trigueros, I. M. (2023). Digital skills and ethical knowledge of teachers with TPACK in higher education. *Contemporary Educational Technology*, *15*(2). <https://files.eric.ed.gov/fulltext/EJ1385486.pdf>
41. Mishra, P., Koehler, M. J., & Henriksen, D. (2011). The seven trans-disciplinary habits of mind: Extending the TPACK framework towards 21st century learning. *Educational Technology*, 22-28.‏ <https://www.researchgate.net/profile/Punya-Mishra-2/publication/267719396_The_7_trans-disciplinary_habits_of_mind_Extending_the_TPACK_framework_towards_21st_Century_Learning/links/547494810cf29afed60f88ca/The-7-trans-disciplinary-habits-of-mind-Extending-the-TPACK-framework-towards-21st-Century-Learning.pdf>
42. Shamsuddinova, S., Heryani, P., & Naval, M. A. (2024). Evolution to revolution: Critical exploration of educators’ perceptions of the impact of Artificial Intelligence (AI) on the teaching and learning process in the GCC region. *International Journal of Educational Research*, *125*, 102326.
43. Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: principles, policy & practice*, *5*(1), 7-74.
44. Mansory, M. (2019). Conversing in L2 English with Saudi Arabic (L1) Children at Home. *Arab World English Journal*, *10*(1), 3-15.
45. Keerthiwansha, N. W. B. S. (2018). Artificial intelligence education (AIEd) in English as a second language (ESL) classroom in Sri Lanka. *Artificial Intelligence*, *6*(1), 31-36**.**
46. Cassidy, C. (2023). Australian universities to return to ‘pen and paper’exams after students caught using AI to write essays. *The Guardian*, *10*.
47. Tapalova, O., & Zhiyenbayeva, N. (2022). Artificial intelligence in education: AIEd for personalised learning pathways. *Electronic Journal of e-Learning*, *20*(5), 639-653. <https://files.eric.ed.gov/fulltext/EJ1373006.pdf>
48. Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology,* *3*(2), 77-101. <https://www.tandfonline.com/doi/pdf/10.1191/1478088706qp063oa?casa_token=b2BcYBLc4H0AAAAA:TFib4fYxWD6lAnUQFWCdy6UUBlTZ6eg0ia3VRuNZWRNVLADMYNq80eU8PyGmvRHfvV89p9ey1rul>
49. Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.‏
50. Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford University Press
51. Rubin, H. J., & Rubin, I. S. (2012). *Qualitative interviewing: The art of hearing data* (3rd ed.). Los Angeles, CA: Sage.
52. Teddlie, C. & A. Tashakkori (2009). *Foundations of mixed methods research*. Thousand Oaks, CA: Sage Publications.