*Original Research Article*

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

.

ABSTRACT

|  |
| --- |
| **Aims:** This study explores Saudi English teachers’ perceptions of the ethical guidelines surrounding the use of AI-generated EFL formative assessments. It also investigates the ethical challenges teachers may encounter when implementing AI in EFL classrooms.**Study Design:** An explanatory sequential mixed-methods design was employed. The study began with quantitative data collection via a questionnaire to examine teachers’ perceptions. This was followed by semi-structured interviews to gain deeper insight into the ethical challenges identified, allowing for elaboration on the initial 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 included 39 Saudi EFL teachers. Of these, 22 were male (57.89%) and 17 were female, ranging in age from 20 to 51 and above, with diverse educational backgrounds and teaching experience. Most participants reported limited use of AI in formative assessment design. Quantitative data were analyzed using SPSS, while qualitative data from interviews with three teachers (two males, one female) were analyzed thematically.**Results:** Descriptive statistics indicated gender-based differences in responses. A t-test (t = -18.15 < 2.042) confirmed a significant difference in one section of the questionnaire, possibly due to variations in experience, perceptions, or priorities. However, the second section showed aligned perceptions across genders.**Conclusion:** The integration of both data sets revealed a consistent theme: while teachers acknowledge AI's potential to enhance formative assessment, they lack the ethical guidance and institutional support needed for its responsible use. Enthusiasm for AI is thus moderated by ethical concerns and practical limitations. |

*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 belong to the pioneers Aiken and Epstein when they wrote an article on Ethical guidelines for AI in education (Latham, Annabel & Goltz, Sean, 2019) 20 years ago. In their paper, they predicted that by 2010, AI would dominate the educational context and teachers’ roles will change. Despite their optimistic perspective towards AI, their cautious tone was loud in the paper especially in regard to ethical utilization of AI in educational settings. Aiken and Epstein (2000) argue that AI has the potential to negatively impact young learners' ethics, aesthetics, physical well-being, emotions, intellect, and social skills. Although this article is considered the kick-start of the discussion of ethics in relation to AI, the focus was mainly on the effect on the students. Teachers’ perspective and their application of AI in their classrooms was not considered. Additionally, it lacks the consideration of the institutions and their role in setting clear principles for the ethical use of AI. It emphasizes the theoretical aspect without mentioning how it could be applied in real-life contexts.

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 ethical issues is privacy which refers to the protection of an individual's data from unauthorized access by others or organizations without explicit consent (Clarke, 1999). In the field of AI ethics, Stahl and Wright (2018) argue that ensuring the protection of personal data within AI systems is a critical issue, given fears that technological companies might gain unauthorized access to this data due to the inherent vulnerabilities in AI systems. Although AI systems would ask their users for their consent to access their private information, various individuals would approve without considering the amount of information that is shared. As for students, data invasion can be exemplified by the personalized learning feature that AI offers since it requires huge data collection. Regardless of its effectiveness to improve the educational system, collecting and storing their data can open a new risk of privacy invasion (Vincent-Lancrin & Van der Vlies, 2020).

Another ethical challenge that may face AI users is the ability to justify and explain AI-generated data. In AI ethics, explainability refers to AI tools’ ability to justify their technical processes and requires humans to comprehend decisions made by them (European Commission, 2019). Goebel et al. (2018) claim that explainability is significant because of the AI’s complicated algorithms, which are designed in such a manner that they conceal their internal rationality and learning methodologies from users. This results in numerous unresolved inquiries regarding the mechanisms and justifications behind the data generated by AI.

Theorodorou, Wortham, and Bryson (2017) explain that the importance of explainability lies in its ability to ensure transparency, provide a measure of reliability, and reveal the decision-making process. In educational settings, teachers need to be able to explain the decisions made by AI to make justifications for their choices inside the classrooms. Additionally, when teachers link the decisions made by AI with learning theories, it assists them in being more confident when using AI. Furthermore, understanding the rationale behind the decisions made by AI tools would lead teachers to trust and accept the generated data and implement them effectively in their teaching practices.

Furthermore, accountability is an additional ethical concern when implementing AI. It is defined by Memarian and Doleck (2023) as the set of protective strategies that make users, owners, and creators of AI charged with the responsibility, meaning that is responsible for what wrongdoing AI generates. From another perspective, Shin, Rasul, and Fotiadis (2022) consider accountability a measurement that is designed to hold the creators of AI systems responsible for the results proposed by their systems. The challenge regarding accountability is answering the question of who is responsible for whatever mistake AI generates. The answer to this question varies according to research. Some researchers believe that AI creators should be blamed for the wrongdoings of AI (Santoni de Sio & Mecacci, 2021). On the other hand, others hold the teachers and the educational stakeholders accountable for not making sure of the accuracy of AI-generated data (Vincent-Lancrin & Van der Vlies, 2020). de Saint Laurent (2018) supports the latter claim. He states that educators cannot blame AI for any inaccurate data. They must make sure of the accuracy of the generated content due to the false claim that AI can work anonymously.

Utilizing AI fairly and equally is another ethical challenge that faces educators. It is referred to as the algorithms that do not create discrimination or bias against certain individuals or groups (Shin, Rasul, & Fotiadis, 2022; Li. Xing, & L). Asr equality, when a system cannot be accessed by the majority, it would not be considered equitable. Stahl and Wright (2018) state that when individuals create algorithms, they also generate data that reflects society’s historical and systemic biases, which eventually lead to algorithmic biases. In the field of education, Khan (2023) illustrates several instances of AI bias. He observes that AI would recommend courses differently based on gender differences. For example, STEM courses were recommended for male students, while arts and humanities were suggested for female students. Additionally, Khan (2023) notes that there is implicit bias in the automated AI grading systems especially in English writing. AI would assign lower grades for students who wrote their essays in non-standard English dialects. The results of this study also highlight the negative impact of these biases on students from marginalized societies as they felt disengaged and frustrated. This study is a serious indicator of the possible threats that AI biases may impose on students’ well-being and teachers’ judgement. The severity of these biases stems from their being embedded and not explicitly shown to educators. As for equality, Nye (2015) states that unequal use of AI tools in societies are not available for every student. He states that there are several challenges that may hinder equal utilization of AI tools in educational contexts including availability of technological devices, availability of electrical power, Internet accessibility, and students’ digital literacy.

Accuracy is another ethical concern that teachers must take into consideration when using AI in their classrooms. Harrer (2023) explains that the AI-generated data may be inaccurate. The reason behind the inaccuracy that AI may generate is that these tools have been trained on a large corpus of raw and unpolished data. Thus, the generated content may be biased and subjective (Sallam, 2023). If there are any biases or inaccuracies in the data being inserted in AI tools it would be shown in the generated responses. In education, students and teachers expressed their concerns regarding the use of AI in their academic practices. In a survey-based investigation, students and teachers stated that despite the human-like responses that AI generates, the accuracy of it cannot be guaranteed, (Chan & Hu, 2023). Similarly, Aguilera-Hermida (2024) discovers that ChatGPT offers inaccurate data that sounds convincing. Also, the in-text references that were provided by this tool were mostly fabricated. Thus, educators need to be aware and not take AI tools for granted.

Furthermore, reliability is considered an ethical issue when discussing AI. Reliable AI-generated content should be objective and consistent. According to Rahayu (2023), since AI tools cannot fully imitate human intelligence, reliability is a huge concern. The reliability of AI results should always be examined by experts. This is because algorithms in the first place are not objective, and they reflect a part of reality (Stahl & Wright, 2018). Moreover, Questioning the reliability of AI-generated data is essential, especially in the field of education. Kim et al. (2024) compare the reliability of AI tools in evaluating students’ placement test essays with that of human raters. The study reveals that ChatGPT’s scoring reliability was moderate to low due to its limitations in detecting topic-specific problems and integrating source text information. The findings emphasize the importance of recognizing the unreliability of AI-generated data and caution against blindly depending on it. The trustworthiness of AI tools is an ethical challenge. Vincent-Lancrin and Van der Vlies (2020) state that AI tools can be considered trustable if they perform properly and fulfill the assigned missions. Notably, trustworthiness is linked to reliability (MI Garage, 2020), explainability (OECD, 2021), and data privacy (Fraunhofer IAIS, 2019). Furthermore, Nguyen et al. (2022) mention that the countries and institutions’ efforts to develop ethical guidelines for the utilization of AI in education are all approaches to trusting AI. In Essence, AI reliability requires objectivity and consistency, whereas trustworthiness focuses on functioning appropriately and the ability to complete assigned tasks effectively.

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 it is not stated clearly, in some empirical studies teachers have expressed their perception towards the ethics of AI when utilizing them in their classroom. Hartono et al. (2023) report in a mixed-methods paper that despite the teachers’ positive perspective towards the utilization of AI, they claim that they lack training and support to use AI effectively in their classrooms. This is an indicator of the ethical principle, which is explainability, since teachers need to practice how to use and explain what AI would generate. In the context of Saudi Arabia, there are little to no research that investigate teachers’ perception of the ethical utilization of AI in educational settings. One study that was conducted across the GCC context on 11 educators’ perspectives towards the implementation of AI. In spite of the small number of participants, they demonstrated their optimistic view toward AI usage in their classrooms. They also highlighted some concerns in regards to the lack of understanding of the AI tools and the lack of structured guidelines that are specific to GCC. Shamsuddinova, Heryani, and Naval (2024) conclude their study by emphasizing the importance of having context-based ethical guidelines for effective implementation of AI tools in education.

Technological, Pedagogical, and Content Knowledge (TPACK) framework include technological content knowledge (TCK), technological pedagogical knowledge (TPK), and technological pedagogical content knowledge (TPACK). Because of its expansive nature and adaptable structure, the framework has been embraced by numerous research studies across various technological devices and pedagogical strategies (Mishra et al., 2010). Ethical guidelines can be linked to TPACK framework from different perspectives. Celik (2023) modifies and extends the original version of TPACK framework to include AI ethics as an additional factor in the scale. The new version is named “Intelligent TPACK framework”. The researcher did not consider AI ethical considerations as an internal component of any of the TPACK framework aspects. The new model integrates AI in every aspect. Moreover, the ethics factor includes solely transparency, accountability, fairness, and inclusiveness. It overlooks other ethical concerns, including explainability, privacy, reliability, and trustworthiness. It is worth mentioning that despite Celik’s focus on the ethical integration of AI among teachers, the instrument did not show an emphasis on it. This can be demonstrated in the number of items each aspect has in the questionnaire. Other factors in the framework, such as intelligent-TK, intelligent-PK, and intelligent-CK each have more than 10 items, whereas ethics includes only four items. Additionally, TPACK has also been implemented to a different focus with reference to ethics. Gómez-Trigueros (2023) investigates teachers’ ethical knowledge when using technology in light of the TPACK theoretical framework. Unlike Celik, Gómez-Trigueros designed a new TPACK version that is catered mainly to investigate the ethics of using technological tools. The questionnaire variables were knowledge of professional ethics (PEK), technological ethical knowledge (TEK), pedagogical technological ethical knowledge (PTEK), disciplinary technology ethical knowledge (EKDT). Notably, the emphasis on ethics can be noticed.

In this current research, the TPACK theoretical framework has been selected to fulfill the goal of exploring Saudi teachers’ perception of the ethical guidelines of AI-generated EFL formative assessments and the ethical challenges that teachers encounter while utilizing the assessments, for several reasons. First, the framework has been proven to be effective to measure teachers’ knowledge of the ethics that are involved with the use of technology, including AI. Celik (2023) and Gómez-Trigueros (2023) studies can be considered as evidence. Also, the framework is flexible and inclusive, allowing for modifications to fit the specific context which is Saudi Arabia. Although the previous studies have not employed the framework for AI-generated EFL formative assessments specifically, the researcher found it manageable to modify the items to suit the intended field, context, and participants. Furthermore, the TPACK framework gives a holistic approach to examine the current study from all aspects including technology, content, and pedagogy. The framework’s components are essential to ensure the effective and ethical utilization of any technological tools, including AI.

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, given the seriousness of the ethical concerns that the utilization of AI may impose on the educational process, specifically in English language classrooms, alongside the importance of English teachers’ perception of these challenges to prevent potential misuse or negative consequences of AI, and the current gap in research on AI ethics in Saudi Arabia, this study investigates Saudi teachers’ perception towards the ethical guidelines of AI-generated EFL formative assessments.

2. methodology

This study employed an explanatory sequential mixed-methods design (Creswell & Creswell, 2018), which begins with quantitative data collection and analysis, followed by qualitative data collection to further explain the quantitative results. This design is widely used in social and educational research due to its capacity to provide both generalizable data and nuanced understanding (Teddlie & Tashakkori, 2009). The primary objective was to investigate Saudi English language teachers' perceptions of AI ethical guidelines and the challenges they encounter when implementing AI-generated formative assessments. Quantitative data were gathered via a structured questionnaire. This was followed by qualitative semi-structured interviews to explore emerging themes in greater depth, particularly on ethical challenges.

Participants consisted of 39 Saudi EFL teachers from the English Language Institute (ELI) at King Abdulaziz University, Jeddah. Participants were selected through convenience sampling based on accessibility and willingness to participate. To ensure contextual relevance, participants were required to be Saudi nationals currently teaching English at the university level. For the qualitative phase, five participants were selected from among those who expressed willingness to be interviewed via the questionnaire. Contact details were obtained through a consent form embedded within the survey.

**2.1 Quantitative data**

The quantitative instrument was a modified 29-item questionnaire adapted from Gómez-Trigueros (2023), originally based on the Technological Pedagogical Content Knowledge (TPACK) framework. Modifications involved aligning the instrument with the study context by focusing exclusively on ethical dimensions of AI use in EFL formative assessment. The scale employed a 5-point Likert response format excluding a neutral option, following Johns’ (2010) recommendations to reduce ambiguous responses in ethically sensitive research. The questionnaire was structured into three sections: (1) demographic data, (2) ethical perceptions in AI-generated formative assessment, and (3) willingness to participate in interviews.

Data was collected online via Google Forms for accessibility and efficiency. All participants were informed of the voluntary nature of their participation and assured of confidentiality. Ethical approval was obtained from the ELI, and consent was documented digitally. Content validity was ensured through expert review by seven PhD scholars in EFL and AI ethics. Cronbach’s Alpha was calculated to assess internal consistency reliability. As shown in Table 1, the instrument exhibited high reliability across all constructs.

**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.

**2.2 Qualitative data**

Semi-structured interviews were developed based on literature (Rubin & Rubin, 2012; Kvale, 2015) and the preliminary quantitative findings. The guide was organized thematically into categories such as experiences with AI, ethical challenges, cultural and religious factors, and institutional support. The guide is included in Appendix A. Interviews were conducted via Zoom, each lasting approximately 30–35 minutes. Participants received the interview guide in advance. All sessions were recorded (with oral consent) and transcribed for analysis. Member checking was employed to ensure accuracy of interpretation (Dörnyei, 2007). Participants were given the opportunity to confirm the authenticity of the transcribed and interpreted data. Data triangulation was achieved by comparing findings from both quantitative and qualitative phases, thereby enhancing validity and reducing researcher bias.

Inductive thematic analysis was used to code and interpret the interview transcripts, following Braun and Clarke’s (2006) six-phase method. NVivo software facilitated coding and theme generation. The process allowed for a systematic examination of recurring ethical concerns and cultural influences on AI use in formative assessment.

This study adhered to strict ethical guidelines. Informed consent was obtained at each stage. Participants were assured of anonymity, and pseudonyms were used in transcripts. Institutional approval was obtained from the ELI, with all documentation secured by research ethics protocols.

3. results and discussion

The study involved 39 Saudi EFL teachers from the English Language Institute (ELI), of whom 57.89% were male and 42.10% were female. Participants varied in age, educational background, and years of experience. The majority held a master’s degree and had between 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 |

**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% |

**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.

**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 |

**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 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**, **data privacy**, 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.

4. Conclusion

This study aimed to explore the challenges that Saudi EFL (English as a Foreign Language) teachers encounter in utilizing ethical guidelines for artificial intelligence (AI) in formative assessment. The primary research question guiding this inquiry was: *What are the challenges that EFL Saudi teachers encounter in utilizing AI ethical guidelines for formative assessments?* Through this investigation, the study sought to identify the gaps in awareness, understanding, and application of ethical principles related to AI use within classroom-based formative assessment practices.

 The central objective was to examine both the practical and ethical implications of integrating AI tools into formative assessment in Saudi EFL contexts. Using a mixed-methods approach, the study collected quantitative data from surveys and qualitative data from interviews and open-ended responses to gain a holistic view of teachers’ experiences, attitudes, and challenges

 The analysis of the data revealed a general openness among Saudi EFL teachers toward using AI in formative assessment; however, it also exposed a lack of sufficient training and institutional guidance in ethical AI practices. Many respondents expressed uncertainty about the ethical implications of data collection, student privacy, and algorithmic decision-making. Furthermore, qualitative findings added depth to this picture, highlighting teachers’ concerns about fairness, cultural appropriateness, and trust in AI tools. Several participants indicated a need for more localized guidelines and context-sensitive frameworks that consider the unique pedagogical and cultural dimensions of Saudi classrooms.

These findings directly addressed the research question by identifying both practical barriers (e.g., lack of training, insufficient infrastructure) and conceptual barriers (e.g., limited understanding of ethical AI principles) that impede ethical implementation.

The integration of quantitative and qualitative findings revealed a consistent pattern: while educators recognize the transformative potential of AI in supporting individualized and timely formative assessment, they simultaneously lack the resources and ethical frameworks necessary for responsible implementation. The convergence of both data types emphasizes that enthusiasm for AI is tempered by significant ethical uncertainty.

Additionally, several themes emerged strongly from the data:

* **Lack of Ethical Awareness:** Many teachers were unaware of existing ethical guidelines or lacked clarity on how to apply them in practice.
* **Institutional Gaps:** There was a perceived disconnect between technological advancement and institutional preparedness.
* **Cultural Sensitivities:** Teachers raised concerns about AI tools not aligning with local values or classroom dynamics.
* **Need for Professional Development:** Teachers expressed a desire for targeted training on ethical AI use.
	1. **Implications for Practice**:

Educators and school leaders should prioritize ethical awareness as part of professional development programs. Training modules that explain data privacy, algorithmic fairness, and responsible AI use in the classroom should be developed and integrated into in-service training. Teachers should also be given access to vetted AI tools that are aligned with pedagogical and cultural standards.

The findings call for the development of national or institutional AI ethics guidelines tailored to educational contexts. Policy makers should engage in collaborative dialogue with educators and technologists to ensure that any AI implementation is grounded in ethical best practices. Policies should also address accountability structures and data governance in educational AI use.

* 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.

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](https://www.diva-portal.org/smash/get/diva2%3A1774035/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](https://unesdoc.unesco.org/ark%3A/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.