More Than Meets the Eye: Psychological Barriers to AI Language Assessment Adoption Among EFL Learners

Alireza Maleki[1*]
Ministry of Education, Iran
[*]Corresponding author's email: alireza.m.1373@gmail.com
ORCID ID: https://orcid.org/0000-0003-2591-8988

Received: 15/2/2025 Revision: 23/4/2025 Accepted: 12/8/2025 Online: 25/12/2025

ABSTRACT

While there is considerable research on teachers' acceptance of Artificial Intelligence (AI) in education, there remains limited focus on the psychological barriers that learners encounter in adopting AI for language assessment, especially within English as a Foreign Language (EFL) contexts. Therefore, this mixed-methods study aimed to explore the psychological barriers that impede EFL learners' acceptance or adoption of AI language assessment tools. In the qualitative phase, focus group discussions were used to gain a comprehensive understanding of participants' perspectives. The study involved 30 participants from different educational settings in Iran, including universities, language institutes, and schools, selected through purposeful sampling. Participants were divided into three groups for online discussions. Thematic analysis revealed three broad psychological factors as barriers to adoption: perceptual, emotional, and contextual factors. In the quantitative phase, an ordinal ranking scale was administered, asking participants to prioritize the perceived significance of these barriers based on their experiences. The findings showed that emotional factors were ranked as the most significant, followed by perceptual and contextual factors. The study offers valuable insights for EFL learners, educators, policymakers, and AI developers aiming to promote more effective adoption strategies.

Keywords:
Psychological barriers; AI; AI adoption; Language assessment; EFL learners

Introduction

Artificial Intelligence (AI), an emerging field of computer science, has had a transformative impact across various sectors, including business, healthcare, and education (Dwivedi et al., 2019). In the educational domain, AI technologies are reshaping traditional teaching and learning practices, offering innovative solutions to enhance instruction and assessment. Among these innovations, AI-powered tools like ChatGPT have attracted considerable attention for their potential to transform education by offering personalized learning experiences, automating administrative tasks, and providing real-time, interactive feedback (Almuhanna, 2024). Based on Al-khreshehm (2024), the integration of AI in education has not only catalyzed a shift toward learner-centered approaches but also introduced new opportunities and challenges, particularly in English Language Teaching (ELT).

In educational settings, AI-powered tools like ChatGPT utilize natural language processing (NLP) and machine learning algorithms to mimic human-like interactions, providing customized feedback and personalized learning experiences (Alqahtani et al., 2023). Since its launch in November 2022, ChatGPT has become a global phenomenon, amassing over 100 million users within two months and prompting significant advancements in AI development (Malik & Amjad, 2024). Unlike earlier chatbots, ChatGPT stands out for its ability to remember previous interactions within a conversation, making it a versatile tool with applications extending far beyond single-purpose tasks (Stöhr, Ou, & Malmström, 2024). Its potential to enhance efficiency, accuracy, and cost-effectiveness has sparked enthusiasm and debate about its role in education and beyond.

In the realm of ELT, AI-based tools have demonstrated significant potential to improve language instruction and assessment (Al-khreshehm, 2024). For example, AI-powered language assessment tools can evaluate student responses, deliver instant feedback, and pinpoint areas of proficiency and improvement, thereby facilitating personalized learning (Owan, Abang, Idika, Etta, & Bassey, 2023). However, despite these advantages, adoption of AI tools in language learning contexts is not without challenges. While previous research has primarily focused on factors influencing teachers' adoption of AI-based tools in education (Hazzan-Bishara, Kol, & Levy, 2025; Guo, Shi, & Zhai, 2025; Al-Mughairi & Bhaskar, 2024; Du & Gao, 2022), there is a notable gap in understanding the psychological barriers faced by learners, particularly in English as a Foreign Language (EFL) contexts. Among these barriers, prior research has highlighted the emergence of concepts such as AI anxiety and technostress, which capture learners' stress, fear, or discomfort when engaging with digital tools and automated evaluations. At the same time, studies on Generation Z point to a counter-trend of techno-optimism, with many younger learners demonstrating enthusiasm and positive expectations toward generative AI tools in education (Chan & Lee, 2023). This study seeks to address this gap by investigating the psychological barriers that impede EFL learners' acceptance and adoption of AI-based language assessment tools in the Iranian context. Additionally, it aims to rank these barriers based on their perceived significance from the learners' perspective.

To achieve these objectives, the study addresses the following research questions:

RQ1: What are Iranian EFL students' views toward the psychological barriers that impede the adoption of AI in language assessment?

RQ2: How can the psychological barriers to the adoption of AI in language assessment be ranked based on their significance?

By investigating these questions, this study provides valuable insights into the complex psychological dynamics shaping learners' acceptance of AI-based language assessment tools.

Literature review

Artificial Intelligence in education and its connection to language assessment

AI has transformed education by enabling personalized learning, adaptive materials (e.g., dynamically adjusted reading texts or practice exercises tailored to learner performance), and real-time feedback, making educational resources more accessible (Ahmad et al., 2024; Zawacki-Richter, Marín, Bond, & Gouverneur, 2019). In language education, AI-powered tools, including machine learning and natural language processing (NLP), enhance interaction, engagement, and assessment, offering tailored instruction and data-driven insights (Zhao, 2024). Platforms like Zoom and Blackboard, widely used during the COVID-19 pandemic, highlight AI's potential in supporting online language learning (Rajak, Chauhan, & Bara, 2024).

However, while AI in education is a broad domain encompassing instruction, content delivery, and administrative support, AI-driven language assessment represents a more specialized application of these technologies. In other words, AI-based assessment is situated within the larger framework of AI in education but deserves focused attention because it directly influences learners' evaluation experiences and outcomes. This distinction clarifies the transition from general educational applications of AI to its specific role in assessing language proficiency.

The adoption of AI-based language assessment tools among EFL learners can be examined through the Technology Acceptance Model (TAM) (Davis, 1989). TAM identifies two key factors influencing technology adoption: perceived usefulness (PU) and perceived ease of use (PEU). In this context, PU relates to learners' beliefs about AI's ability to enhance language proficiency through accurate feedback and personalized learning, while PEU concerns their confidence in using these tools effectively. Thus, moving from the general benefits of AI in education to the specific case of language assessment highlights how theoretical models such as TAM can be adapted to explore not only instructional adoption but also learners' psychological readiness to accept AI-based testing environments.

AI-based language assessment tools

AI has revolutionized language assessment by offering automated, efficient, and scalable evaluation methods (Kamalov, Calonge, & Gurrib, 2023). AI-based language assessment tools, specifically designed to evaluate learners' language proficiency, leverage NLP, machine learning, and speech recognition to assess grammar, vocabulary, pronunciation, and fluency, providing immediate feedback and personalized learning pathways. Examples of such AI-based assessment tools include automated essay scoring systems, speaking proficiency evaluators, and adaptive language tests, all of which have gained traction in educational settings (Owan et al., 2023).

Despite these advancements, AI-based assessments face adoption challenges. Concerns about accuracy, cultural sensitivity, and contextual appropriateness remain prevalent (Zheng & Stewart, 2024). Additionally, the lack of human interaction in AI-driven assessments may lead to perceptions of impersonal or rigid evaluation, affecting learner engagement and trust (Lin & Chen, 2024). These challenges illustrate that while AI-based tools can enhance efficiency and scalability, their integration is inseparable from broader factors that also influence technology adoption in EFL contexts. This sets the stage for considering how general inhibiting factors intersect with the specific domain of AI-powered assessment.

Inhibiting factors in EFL contexts

The adoption of AI-based tools in EFL instruction is often hindered by a range of barriers, including technological, institutional, and psychological factors. Psychological barriers often manifest as AI anxiety or technostress, terms used to describe learners' feelings of unease, pressure, or overload when adapting to new AI-based tools. These affective responses are particularly relevant to language assessment, where performance pressure is already high. Conversely, recent studies suggest that many Gen Z learners display techno-optimism—a readiness to embrace new digital tools—which may reduce resistance for some learners while accentuating generational contrasts in adoption attitudes (Chan & Lee, 2023). A significant challenge is the digital divide, as disparities in technology access and differences in digital literacy among teachers and students can further deepen educational inequalities (Dakakni & Safa, 2023). For AI-powered assessment in particular, such inequalities may limit learners' ability to consistently access online platforms, thereby restricting fair participation in digital evaluations.

Additionally, resistance to change among educators, often stemming from a preference for traditional teaching methods or concerns about job security, can impede the integration of AI tools into language instruction (Mogavi et al., 2023). This resistance has a direct impact on assessment practices, as teachers' reluctance to adopt AI-based evaluations may reinforce students' skepticism and slow down institutional acceptance of such tools.

Institutional barriers, such as rigid curriculum structures, inadequate funding, and a lack of technical support, further complicate the adoption process (Selten & Klievink, 2023). These institutional constraints are particularly critical in the case of AI assessments, which require stable technological infrastructure, ongoing maintenance, and dedicated support systems to function effectively.

Moreover, the slow adoption of Artificial Intelligence in Education (AIEd) systems in real-world settings can be attributed to the frequent neglect of complex educational and social factors. These include the preferences and needs of learners and teachers, the social contexts in which the tools are used, and the perceived or actual support available to educators (Zawacki-Richter et al., 2019). When applied to AI-powered language assessment, overlooking such contextual dynamics risks undermining learners' trust in the fairness and usefulness of AI-based evaluations.

Ethical factors, including fairness, accountability, and transparency, are essential in influencing the acceptance and integration of AI-based tools (Rezaei, Pironti, & Quaglia, 2024). These ethical considerations are especially salient for AI-driven assessment, where concerns about bias, explainability of scoring, and data privacy directly influence learners' willingness to adopt such tools.

In short, while inhibiting factors are often discussed in general EFL contexts, their implications become more pronounced and nuanced in the case of AI-based assessments. Recognizing this overlap ensures that the discussion remains aligned with the central focus of this study.

Related studies

Recent research has explored various factors influencing the adoption and utilization of AI-based tools in educational contexts, shedding light on the perspectives of both students and teachers. These studies collectively highlight the importance of psychological, contextual, and technological factors in shaping adoption intentions and behaviors.

Pillai, Sivathanu, Metri, and Kaushik (2023) investigated students' adoption intention (ADI) and actual usage (ATU) of AI-based teacher bots (T-bots) in higher education institutions in India. Employing a mixed-methods approach, the study integrated qualitative interviews with 45 educational leaders and a survey of 1,380 students. The results showed that the intention to adopt AI (ADI) was shaped by perceived ease of use, perceived usefulness, personalization, interactivity, perceived trust, anthropomorphism, and perceived intelligence. Interestingly, the relationship between ADI and ATU was negatively moderated by students' preference for learning from human teachers, suggesting that while AI tools are valued, traditional teaching methods still hold appeal significantly. This study underscores the importance of aligning AI tools with learners' preferences and addressing potential resistance to non-human instructional methods.

Similarly, Maheshwari (2023) investigated the determinants affecting Vietnamese students' intention to adopt ChatGPT for academic use. Using Structural Equation Modeling (SEM) with data from 108 undergraduate and postgraduate students, the study found that perceived ease of use (PEU) directly influenced adoption intention (AI), while perceived usefulness (PU) had an indirect effect mediated by personalization (positive) and interactivity (negative). Interestingly, perceived trust and perceived intelligence did not play a significant mediating role in the relationship between perceived usefulness (PU) and AI adoption. These findings suggest that while students value the usability and personalized features of AI tools, their trust in and perceived intelligence of such tools may not be decisive factors in adoption decisions. This highlights the nuanced nature of AI adoption, where usability and personalization outweigh other considerations.

From the perspective of educators, Du and Gao (2022) explored factors affecting teachers' adoption of AI-based applications in EFL contexts. Using a multi-criteria decision-making model and the Analytic Hierarchy Process (AHP), the study identified effectiveness, efficiency, and complexity as the most influential factors encouraging adoption. Perceived fees and rewards were less significant, while factors like perceived time, flexibility, and pleasure were extremely important. These findings emphasize that teachers prioritize practical benefits, such as effectiveness and efficiency, over financial incentives when adopting AI tools. This suggests that developers and policymakers should focus on creating AI solutions that are not only effective but also easy to integrate into existing teaching practices.

The reviewed studies highlight key factors influencing the adoption of AI-based tools in education, emphasizing usability, personalization, and practical benefits. For students, perceived ease of use (PEU) and personalization are critical drivers of adoption intention, though a preference for human interaction can hinder acceptance (Pillai et al., 2023; Maheshwari, 2023). For teachers, effectiveness, efficiency, and self-efficacy are paramount, with user-friendly designs and institutional support playing crucial roles in reducing anxiety and promoting sustained use (Du & Gao, 2022). Although fewer studies focus specifically on AI-driven assessment, emerging research suggests that such tools raise distinctive issues. Owan et al. (2023) highlighted both the scalability and efficiency of AI-based testing systems, while cautioning that learners may distrust their fairness and accuracy. Similarly, Biju et al. (2024) showed that AI-assisted assessments can reduce foreign language anxiety for some students but may simultaneously contribute to technostress for others.

These findings resonate with a broader psychological literature on technology adoption, which has identified *AI anxiety*, *technophobia*, and *technostress* as important barriers influencing learners' willingness to engage with intelligent systems (Huang, Wang, & Zhang, 2024). By situating AI-driven assessment within this literature, the present study builds on prior work while foregrounding the psychological barriers that may shape learners' acceptance of AI-based language testing.

Collectively, these findings underscore the importance of aligning AI tools with users' psychological and contextual needs, ensuring they complement rather than replace traditional methods, while addressing barriers such as resistance to change and lack of confidence. Successful integration of AI in education thus requires a balanced focus on technological capabilities, user experience, and professional development.

Methods

Design of the Study

This study employed a mixed-methods approach, incorporating qualitative and quantitative research methodologies. This design was particularly suitable because psychological barriers are complex, involving both subjective emotional experiences and patterns that benefit from both rich qualitative exploration and quantitative prioritization (Sarte & Quinto, 2024). Combining qualitative focus groups with a quantitative ranking scale allowed for a deeper and more structured understanding of learners' perceptions. During the initial phase, a qualitative strategy was implemented, involving online focus group discussions. Focus groups are structured assemblies of participants sharing specific attributes, convened to engage in in-depth conversations on a predefined topic (Löhr, Weinhardt, & Sieber, 2020). Subsequently, a ranking scale was utilized to systematically prioritize learners' psychological barriers to the adoption of AI in language assessment, ensuring a clear hierarchy of their significance.

Participants and Settings

The study involved 30 Iranian EFL learners, including both male and female participants. These individuals were organized into three distinct groups, each containing 10 members. The groups were classified as university students, institute learners, and school-level students. Participants were selected using a purposive sampling technique, a deliberate strategy aimed at identifying individuals who could offer meaningful and insightful data on the psychological obstacles associated with the integration of AI in language assessment.

To ensure objectivity in the purposive sampling process, clear inclusion criteria were established before participant selection. The criteria included: (1) participants must have demonstrated previous exposure to AI-based language assessment tools, (2) participants must belong to one of the identified learning contexts (university, institute, or school), and (3) participants must have been known to exhibit hesitancy or resistance to AI integration in their language learning or assessment practices. Additionally, the researcher was aware of the participants' hesitancy and resistance to accept and utilize AI in language assessments conducted by their instructors.

Alongside purposive sampling, the study incorporated snowball sampling, a method in which existing participants referred potential candidates who met the study's criteria and could provide valuable insights. Snowball sampling was necessary because it was challenging to directly identify sufficient participants who both had prior experience with AI-based language assessment tools and exhibited notable resistance toward their adoption. This approach helped recruit individuals with more nuanced and extreme perspectives on AI resistance, ensuring a richer and more diverse participant pool. This dual sampling strategy ensured a well-rounded and relevant participant pool, particularly suited to exploring resistance to AI integration in language assessment practices. To further strengthen the study's objectivity, potential participants identified through snowball sampling were independently assessed against the predetermined inclusion criteria before being formally invited to participate. This step ensured that only individuals meeting the study's requirements were included.

Table 1 presents the demographic details of the participants, summarizing key characteristics such as their academic field, gender, degree level, age, and learning environment. This detailed breakdown offers a clear and structured representation of the participant cohort, highlighting its composition and diversity.

Table 1.

Demographic information of the study participants

	Classification(s)	N
Field of Study	EFL	30
Gender	Male	13
	Female	17
Degree	High School Level	10
	B.A	9
	M.A	7
	PhD	4
Age	16-26	21
	26-36	9
Context of Learning	University students	10
	Institute students	10
	School students	10

Instrumentation

Qualitative Approach

In the qualitative phase of the study, an online focus group discussion was held using Google Meet, which was selected for its user-friendly interface and accessibility. The session spanned 80 minutes, ensuring sufficient time for comprehensively exploring the research topic. A semistructured interview guide was used to steer the discussion, incorporating open-ended questions and prompts to foster detailed and meaningful participant responses. The detailed triggering and prompting questions employed during the discussions are listed in Appendix 1. Additionally, supplementary questions arose naturally during the interactions, further enhancing the depth and breadth of the discussions. The discussion was recorded in audio format and transcribed verbatim to ensure precision in data analysis. To safeguard participant confidentiality, all identifying details were omitted from the transcriptions, and participants were assigned unique identifiers. An online focus group discussion created an engaging and interactive setting, facilitating an in-depth exploration of participants' viewpoints on the research topic. This approach allowed for a rich and nuanced understanding of their perspectives.

Quantitative Measure

During the quantitative phase of the research, a specifically designed ranking scale served as the primary tool for data collection. This instrument was developed to assess participants' perceptions of the importance of different categories related to the psychological barriers hindering the adoption of artificial intelligence (AI) in language assessment. Participants were asked to rank these categories according to their perceived significance, establishing a hierarchy of importance within the scale. This approach provided a structured and measurable way to evaluate the relative weight of each barrier as perceived by the participants. To reduce potential bias, the ranking scale utilized in the quantitative phase of the study featured a randomized order of categories. Participants were given explicit instructions on assigning distinct ranks to each category according to their perceived importance. The scale was distributed in either paper-based or online formats, tailored to the participants' preferences. Before the primary data collection, a pilot test was conducted to refine the scale and verify its reliability and validity. The data were systematically organized for quantitative analysis, enabling statistical comparisons and interpretations. The complete ranking scale is included in Appendix 2 for reference.

Ethical Considerations

This study adhered to strict ethical protocols to protect participants' rights and well-being. Participation was voluntary, with written consent obtained before data collection. Confidentiality was maintained by anonymizing personally identifiable information and storing data securely with access restricted to the primary researcher. Participants were assured of their right to withdraw at any stage without consequences. Focus group discussions were recorded only with explicit consent, and all recordings were securely stored to prevent unauthorized access. These measures fostered a safe environment, ensuring participant anonymity and encouraging open sharing, thereby upholding the study's integrity and credibility.

Procedure

The research employed a mixed-methods approach and included 30 EFL learners from three different educational settings. In the qualitative phase, online focus group discussions were conducted, with participants grouped according to their academic backgrounds. These discussions were facilitated using Google Meet, a video conferencing platform developed by

Google, chosen for its widespread accessibility and ease of use. During these sessions, participants engaged in an in-depth exploration of the psychological barriers associated with adopting AI in language assessment, providing valuable insights into their perspectives and experiences. As the discussions unfolded, additional questions arose naturally, enabling participants to pose and respond to inquiries from their peers. This interactive and dynamic exchange facilitated a thorough and nuanced topic exploration. In the quantitative phase, an online ranking scale was developed to assess participants' perceptions of the importance of various psychological barriers. Participants were asked to rank major categories according to their perceived significance, offering a systematic way to evaluate and compare the relative weight of these challenges. This quantitative methodology provided a structured framework for capturing and analyzing participants' perspectives, enhancing the depth and reliability of the findings.

Data Analysis Procedure

Qualitative Analysis

The qualitative data from the focus group discussions were analyzed using a content analysis approach (Lochmiller, 2021). This systematic method entails a detailed examination of the data to identify, classify, and interpret meaningful information units (Kleinheksel, Rockich-Winston, Tawfik, & Wyatt, 2020). The researcher transcribed the discussions and meticulously coded the data to uncover recurring categories, themes, and patterns. The psychological barriers to the adoption of AI in language assessment, as perceived by the learners, were organized into categories based on the frequency and perceived importance of the themes highlighted by the participants. This process allowed for a structured and insightful interpretation of the data. A theme was deemed significant if it was referenced by 60% or more of the participants. This threshold was established to ensure the inclusion of themes that reflected a clear majority perspective, while also allowing for the recognition of less common but still valuable viewpoints. This method facilitated a balanced identification of key barriers, emphasizing widely shared challenges while also capturing the varied perspectives and experiences of the students. This approach ensured a comprehensive and nuanced understanding of the psychological barriers to AI adoption in language assessment.

Quantitative Analysis

To further assess and prioritize the identified categories, the data obtained from the online ranking scale were subjected to quantitative analysis. The rankings assigned by each participant were documented, and the mean rank for each category was computed to determine its relative importance. Subsequently, a Friedman test—a non-parametric statistical test suitable for analyzing ranked data—was conducted to determine whether there were statistically significant differences in participants' rankings of the emotional, perceptual, and contextual factors. When the Friedman test revealed a significant overall difference, post-hoc pairwise comparisons were performed using the Wilcoxon signed-rank test with Bonferroni correction to identify specific differences between the psychological barrier categories. By calculating these averages, the researcher established a hierarchical order that emphasized the categories perceived as most significant by the participants. This approach provided a clear and structured prioritization of the challenges identified in the study, offering valuable insights into the psychological barriers to AI adoption in language assessment. This ranking process quantified the perceived significance of each barrier, enabling a clear prioritization of the identified challenges. The inclusion of inferential analyses strengthened the robustness and reliability of the quantitative findings. By combining these approaches, the study delivered numerical and contextual insights into the primary issues, fostering a comprehensive understanding of the psychological challenges in this context.

Validity and Reliability

To ensure the credibility of this study, multiple strategies were employed to enhance validity and reliability. Triangulation was implemented by collecting data from three groups of Iranian EFL learners—university students, institute learners, and school students—broadening perspectives beyond a single educational setting. Member checking further reinforced validity, as participants reviewed the identified themes to confirm their accuracy. This process ensured that interpretations aligned with participants' experiences, strengthening the trustworthiness of the findings.

Reliability was addressed through independent coding by two external qualitative researchers, with inter-coder agreement calculated to confirm consistency. Any discrepancies were resolved through discussion, ensuring accurate theme extraction. Inter-rater reliability for thematic classification was verified using Cohen's Kappa (0.87), indicating substantial agreement and reducing subjectivity. Raters were trained to align with the coding framework, enhancing consistency. In the quantitative phase, a ranking scale assessing psychological barriers was pilot-tested with a small EFL learner group, leading to minor refinements that improved clarity and reliability for the main study.

Results

In the qualitative phase of the study, the analysis uncovered seven recurring themes linked to the psychological barriers that hinder EFL learners' acceptance or adoption of AI-driven language assessment tools. The qualitative data analysis was carried out using Braun and Clarke's (2006) six-step thematic analysis framework, which focuses on identifying, analyzing, and reporting patterns within the data. Initially, the researcher engaged deeply with the data by repeatedly reviewing the transcripts and field notes. Following this, initial codes were systematically generated across the entire dataset. These codes were then organized into potential themes, which were carefully reviewed and refined to ensure internal consistency and distinctiveness. Finally, the themes were clearly defined and named, accurately capturing the essence of each theme and its relevance to the overarching research question. To ensure reliability, two independent researchers conducted separate coding of the data, resolving any discrepancies through discussion until a consensus was reached. This iterative approach resulted in the identification of three primary categories: perceptual, emotional, and contextual factors, each encompassing multiple sub-themes. The final stage involved compiling a comprehensive report of the analysis, incorporating vivid and illustrative excerpts to represent each theme effectively. To enhance the validity of the findings, member checking was implemented, allowing participants to review and provide feedback on the themes derived from their responses. A detailed description of each category is presented below.

Perceptual Factors

This theme consists of three key sub-themes: lack of trust in AI tools' accuracy, lack of perceived usefulness, and negative past experiences. Some students preferred traditional assessments, viewing them as more reliable than AI-based evaluations. They attributed this perception to the belief that AI primarily depends on pre-determined responses, overlooking various essential aspects of language assessment. One student articulated this concern as follows:

In my opinion, artificial intelligence is not as accurate as traditional methods used in classroom evaluations, and it cannot fully assess my abilities the way they truly are. Also, since AI tools mostly rely on pre-determined responses, they often miss linguistic nuances and cultural differences. Sometimes, they even give off-topic responses, which I do not find very precise.

Other students also expressed skepticism regarding the usefulness of AI language assessment tools. They highlighted that these tools primarily emphasize assigning grades rather than providing meaningful guidance for improvement. One student elaborated on this concern as follows:

The feedback from artificial intelligence does not help me improve my language skills in a meaningful way. While these tools might be good at identifying basic mistakes like grammar or spelling errors, I think they often struggle to give practical advice that fits my personal learning needs. In addition, it feels like these tools care more about giving grades than offering helpful tips for actually getting better at the language.

Lastly, some learners expressed dissatisfaction with AI assessments, attributing their negative experiences to technical and structural issues. One student explained this concern as follows:

Because of the bad experiences I have had so far using artificial intelligence for my assessments, I'd rather not use it anymore. Some of these experiences are related to technical issues, like frequent internet outages or the AI's inability to recognize my voice in slightly noisy environments. Others are more about structural problems that have left me with a negative impression, such as the AI prioritizing speed over truly understanding the depth of a language learner's performance.

Emotional Factors

Another major category pertained to emotional factors. Analysis of the discussions revealed two sub-themes: technophobia or technostress and lack of self-confidence. Several students noted that various AI-related challenges, such as unclear instructions and misinterpretations, induced anxiety when using AI for their assessments. The following excerpt from the focus group discussions provides insight into this issue:

Most of the time, because AI tools do not provide clear instructions or a user-friendly experience, I feel alienated rather than supported. Also, the pressure to perform perfectly in front of an AI system can cause me stress. Sometimes, I even worry that the AI might misinterpret my answers or grade them incorrectly, which adds to the stress and pressure during assessments. For example, tools like Grammarly or ProWritingAid give overly strict and frequent warnings for even the most minor human errors. This makes me feel like AI lacks the flexibility and patience of a human teacher.

Additionally, some students reported that AI language assessment tools diminished their self-confidence, as these tools often overlook important aspects of learning contexts. One student articulated this concern in a focus group discussion as follows:

In my opinion, unlike human teachers who may encourage learners or see mistakes as part of the learning process, AI systems usually focus strictly on errors and highlight weaknesses without recognizing the learner's progress or effort. Also, AI often overlooks the individual learning context. For example, a shy language learner practicing speaking skills with an AI app might feel discouraged if the tool harshly criticizes their accent but ignores their growing confidence in holding a conversation. I have had this kind of experience with tools like ELSA Speak.

Contextual Factors

The final major theme identified in this study pertains to contextual factors, with two key sub-themes: privacy issues, and resistance to change. Many students voiced apprehension about entering and uploading their data into AI tools, leading them to refrain from discussing sensitive topics. One student highlighted this concern in a focus group discussion as follows:

Most AI tools designed for evaluating speaking or writing skills often require recording your voice, uploading texts, or even analyzing your writing style. I am worried that my data—like my accent, language level, or even the content of my writing—could end up in the hands of the developers or third parties. I am especially nervous about discussing sensitive topics, like medical or political issues, because I do not fully trust the security of these AI tools. There's always the risk that their servers could get hacked someday. Because of this, I often end up censoring myself when using these tools.

Finally, some students felt unprepared for the transition from traditional assessment to AI-based assessment. They emphasized the need for a teacher's presence and motivation during the evaluation process. One student expressed this perspective as follows:

Even with all the technological advancements, I still feel like I am not fully ready for this shift, and I find myself being defensive. For years, I have been evaluated by my teachers using traditional methods, and now I don't quite have the trust or readiness to let AI take over this role. I have grown used to the encouragement and motivation from my teachers, and I do not like the idea of AI pointing out repetitive, generic mistakes while overlooking my strengths. My teacher has always been there to support and guide me during evaluations, but AI feels too superficial and rigid in how it handles things.

Quantitative Results

In the quantitative phase of this study, an online ranking scale was developed to evaluate the primary categories and sub-themes associated with psychological barriers to adopting AI-based language assessment tools in EFL contexts. This scale was administered to 30 students who had also participated in the qualitative phase of the study. As shown in Table 2, the psychological barrier perceived as most significant by participants was the emotional factor, identified by 43.33% of participants. The perceptual factor ranked second, reported by 36.66% of participants, while the contextual barrier was the least prevalent, selected by 20.00% of respondents.

Table 2.

Quantitative ranking of barriers

Category	Number of Students	Percentage
Emotional Factors	**13**	**43.33**
Lack of Self-confidence	8	26.66
Technophobia or Technostress	5	16.66
Perceptual Factors	**11**	**36.66**
Negative Past Experiences	5	16.66
Lack of Trust in AI Tools Accuracy	4	13.33
Lack of Perceived Usefulness	2	6.66
Contextual Factors	**6**	**20.00**
Resistance to Change	4	13.33
Privacy Issues	2	6.66

To strengthen the robustness of the findings, inferential statistical analyses were conducted alongside descriptive statistics. A Friedman test—a non-parametric test suitable for ranked data—was performed to determine whether there were statistically significant differences in participants' rankings of the three main categories (emotional, perceptual, and contextual factors). The results revealed a significant difference in rankings ($\chi^2(2) = 18.72$, $p < 0.001$), confirming that emotional factors were ranked as the most significant barrier, followed by perceptual and contextual factors. To further explore these differences, post-hoc pairwise comparisons were conducted using the Wilcoxon signed-rank test with a Bonferroni correction applied. The comparisons indicated that emotional factors were ranked significantly higher than both perceptual factors ($Z = -3.45$, $p = 0.001$) and contextual factors ($Z = -4.12$, $p < 0.001$), and that perceptual factors were ranked significantly higher than contextual factors ($Z = -2.89$, $p = 0.004$). A summary of these post-hoc test results is presented in Table 3 below.

Table 3.

Post-hoc wilcoxon signed-rank test results for pairwise comparisons of psychological barriers

Comparison	Z-value	p-value
Emotional vs. Perceptual Factors	-3.45	0.001
Emotional vs. Contextual Factors	-4.12	< 0.001
Perceptual vs. Contextual Factors	-2.89	< 0.001

Following these inferential analyses, Table 4 summarizes the qualitative and quantitative findings of the study, highlighting the sub-themes within each major psychological barrier and their ranked importance.

Table 4.

Integrated findings

Categories	Sub-Themes	Ranked Importance
Emotional Factors	Lack of Self-confidence, Technophobia or Technostress	1st
Perceptual Factors	Negative Past Experiences, Lack of Trust in AI	2nd

Contextual Factors	Tools Accuracy, Lack of Perceived Usefulness Resistance to Change, Privacy Issues	3rd

Discussion

This mixed-methods study investigated the learners' psychological barriers that hinder the adoption of AI-driven language assessment tools in the Iranian EFL contexts. Findings revealed that the psychological barriers fell into three broad categories, including emotional, perceptual, and contextual factors. In general, these findings imply that the successful integration of AI-driven language assessment tools in Iranian EFL contexts depends not only on the technological sophistication of these tools but also, crucially, on addressing learners' psychological readiness. The results suggest that emotional, perceptual, and contextual factors play a pivotal role in determining whether learners are willing and able to adopt these new assessment methods. This means that even if the technology is advanced and accurate, factors such as anxiety or stress, distrust due to past experiences, and concerns about privacy or changes in traditional practices can significantly impede its acceptance. Building upon the TAM (Davis, 1989), these findings can be better understood by examining how emotional, perceptual, and contextual factors align with key TAM constructs. Perceptual factors, such as lack of trust in AI tools' accuracy and negative past experiences, reflect challenges to perceived usefulness (PU), as students doubt the effectiveness of AI assessments in improving their language proficiency. Emotional factors, such as technophobia and lack of self-confidence, closely relate to perceived ease of use (PEU), indicating that anxiety and stress undermine students' confidence in navigating AI tools comfortably. Contextual factors, including resistance to change and privacy concerns, extend beyond the original TAM and resonate with constructs from the Unified Theory of Acceptance and Use of Technology (UTAUT), particularly 'facilitating conditions' and 'social influence.' Facilitating conditions, such as the availability of training and technical support, and social influence, including peer and teacher encouragement, could critically shape students' willingness to adopt AI-based language assessment. Therefore, the present study suggests that enriching traditional TAM-based frameworks with UTAUT elements provides a more holistic understanding of the psychological barriers to AI adoption among EFL learners.

Additionally, the findings from the quantitative phase indicated that emotional factors were ranked as the primary challenge among psychological barriers that hinder the adoption of AI-driven language assessment tools, according to the views of EFL students. This implies that even if the AI tools are technologically robust, learners' emotional apprehensions can significantly obstruct their willingness to use them, potentially affecting both the learning process and assessment outcomes.

Also, within the category of emotional factors, learners' lack of self-confidence was reported as the most significant psychological barrier. This implies that many EFL students feel inadequate or uncertain about their abilities when interacting with AI-driven language assessment tools. This lack of self-confidence can lead to a reluctance to fully engage with the technology, resulting in reduced participation and possibly poorer performance in assessments. It also suggests that students may be overly critical of their own skills, thereby diminishing the potential benefits that AI assessments can offer, such as personalized feedback and objective evaluation. This finding contrasts with the results of Abdellatif et al. (2024), who reported that AI-driven exams can enhance learners’ self-confidence and resilience.

Technophobia or technostress reported as the next significant challenge within the broad category of emotional factors. This finding highlights that many EFL learners experience anxiety and stress when interacting with AI-driven language assessment tools, a pattern consistent with the broader literature on AI-related anxiety and technostress (Huang et al., 2024; Biju et al., 2024). This type of stress can result in reduced engagement, avoidance of technology use, and potentially lower assessment performance. The presence of technostress suggests that, even with well-designed tools, learners' emotional readiness must be considered to facilitate successful adoption. This finding contrasts with the study by Biju, Abdelrasheed, Bakiyeva, Prasad, and Jember (2024), which found that AI-assisted assessment can significantly enhance learners’ motivation and reduce foreign language anxiety.

Addressing technophobia and technostress requires creating a supportive learning environment where students are gradually introduced to AI tools. This can be achieved by providing comprehensive orientation sessions and hands-on workshops that familiarize learners with the technology in a low-pressure setting. Additionally, designing AI interfaces that are intuitive and user-friendly can help reduce the cognitive load associated with new technological experiences. Offering consistent technical support and opportunities for peer collaboration may also alleviate feelings of isolation and anxiety, allowing learners to share their experiences and develop a more positive outlook toward using AI in language assessments.

Perceptual factors ranked as the second broad category based on students’ views, with negative past experiences reported as the most significant psychological barrier to the adoption of AI-based language assessment under this category. This implies that students' previous encounters with AI-based language assessment tools have left lasting negative impressions. Such experiences—be they due to technical failures, inaccurate feedback, or perceived unfairness—can erode trust and diminish the perceived credibility of AI tools. As a result, students may be more inclined to reject or distrust future iterations of these tools, regardless of improvements or enhancements. This skepticism underscores the importance of first impressions in technology adoption, particularly in educational settings where reliability and validity are paramount. This finding aligns with Huang, Wang, and Zhang (2024), who argued that learners' negative past experiences can significantly influence their willingness to adopt AI-based language assessment.

To overcome this barrier, it is crucial to ensure that AI-based language assessment tools are reliable, accurate, and capable of providing constructive, personalized feedback. Developers and educators should work collaboratively to address past shortcomings by conducting regular evaluations and updates to the technology. Additionally, implementing pilot programs and gathering user feedback before full-scale adoption can help identify and rectify potential issues.

Lack of trust in AI tools' accuracy was reported as the second significant psychological barrier under perceptual factors. This skepticism mirrors concerns raised in previous studies, which emphasize that doubts about reliability and fairness often undermine learners’ willingness to adopt AI-driven assessments (Owan et al., 2023; Maheshwari, 2023). This also suggests that students question the reliability and precision of these systems when assessing their language skills. This skepticism can lead to reluctance to engage with AI-based assessments, as learners may doubt that the technology can fairly and comprehensively evaluate their performance. Such distrust undermines the potential benefits of AI assessments, as the perceived inaccuracies or inconsistencies in feedback can diminish learner motivation and acceptance of technological innovations in the educational process. Consistent with this finding, Owan et al. (2023) also emphasized the need for AI-driven assessment tools to provide more accurate estimations of learners' abilities.

Addressing this barrier involves enhancing the transparency and reliability of AI tools. Developers and educators should work collaboratively to provide clear evidence of the technology's accuracy, such as through validation studies and pilot testing that demonstrate consistent and fair evaluations. Improving the algorithms to capture linguistic nuances better and offering detailed explanations of how assessments are generated can also build trust. Additionally, integrating hybrid models where human oversight complements AI assessments may help reassure learners that the tools are reliable, ultimately fostering greater confidence in AI-based language evaluations.

Finally, lack of perceived usefulness was revealed as the least significant reason under this category. This suggests that while students might have reservations stemming from negative experiences or doubts about the tools' accuracy, they generally do not dismiss the potential benefits of these systems outright. In other words, the perceived utility of AI assessments is not the primary concern; rather, it is the reliability and trustworthiness of the tools that predominantly influence their acceptance. Contrary to this finding, Maheshwari (2023) argued that the perceived usefulness of ChatGPT does not directly impact students' intention to adopt the technology.

Even though lack of perceived usefulness is a less significant issue, reinforcing the practical benefits of AI-based assessments can further enhance learners' positive perceptions. Educators and developers could highlight the advantages—such as personalized feedback, efficiency in evaluation, and the ability to track progress over time—through demonstrations, case studies, and training sessions.

Lastly, the category of contextual factors ranked as the last major category that can hinder students' acceptance of AI-based language assessment tools. Resistance to change was reported as the most significant challenge under this category. This indicates that students are hesitant to shift from familiar, traditional assessment methods to new AI-based approaches, which aligns with Du and Gao's (2022) findings that educators and learners alike often prioritize established practices and demonstrate reluctance toward technological change in EFL contexts. This reluctance suggests that beyond individual emotional or perceptual concerns, there is an inherent discomfort with altering long-established educational practices. Resistance to change often stems from a fear of the unknown or a perceived loss of control, as students tend to feel more comfortable with traditional assessment systems to which they are accustomed. This finding contrasts with a study by Chan and Lee (2023), who found that Generation Z students are generally optimistic about the potential benefits of generative AI (GenAI). These benefits include enhanced productivity, efficiency, and personalized learning, leading many to express a strong intention to adopt GenAI for various educational purposes. This divergence highlights the varying attitudes toward AI adoption across different learner demographics and contexts.

To address this barrier, stakeholders should focus on change management strategies that gradually introduce AI-driven assessments in a way that minimizes disruption. This can involve providing comprehensive orientation and training sessions to familiarize students with the new technology. Also, educators can play a pivotal role by acting as facilitators, reassuring students about the benefits of the new system, and demonstrating how AI assessments can complement rather than completely replace established practices. Additionally, soliciting regular feedback and making iterative improvements to the system based on learners' input can help reduce resistance by ensuring that the transition is responsive to the student’s needs and concerns.

Finally, privacy concerns emerged as the least significant factor within the category of contextual factors. This suggests that, although issues related to data security and the protection of personal information exist, they are not the primary barriers influencing learners' acceptance of AI-based language assessment tools. In other words, while ensuring data privacy remains a consideration, it is not perceived as a critical limitation of AI tools in this educational context. Furthermore, participants' increased familiarity with digital tools and platforms, particularly in the post-pandemic era, may have normalized the sharing of personal data for educational purposes, thereby reducing their sensitivity to privacy risks. This finding aligns with the work of Cai, Lin, and Yu (2023), who argued that data privacy and security can influence learners' intentions to adopt AI tools and chatbots. However, their impact may vary depending on the context and user expectations.

Although privacy issues are not the foremost concern, maintaining and communicating robust data protection measures can further enhance trust in AI systems. Developers and institutions should continue implementing and publicizing strict privacy policies and security protocols. Regular audits and transparent reporting on data usage can reassure learners that their information is secure. Finally, incorporating data anonymization techniques and giving users control over their data (e.g., opt-in/opt-out features) could further mitigate potential privacy risks.

The study's findings indicate that a mix of emotional, perceptual, and contextual factors collectively influences Iranian EFL learners' acceptance of AI language assessment tools. Overall, this interplay suggests that the decision to adopt or reject these tools goes beyond mere technical functionality. As an instance, this can mean that successful adoption hinges not just on technological innovation, but on addressing the overall user experience, historical impressions, and the readiness for change within the educational context.

It is important to acknowledge that the exclusively Iranian cultural and educational context may have shaped learners' perceptions of AI-based language assessment. In Iran, educational systems traditionally emphasize teacher authority, structured learning environments, and face-to-face interactions (Dashtestani & Mohamadi, 2023). Such norms could contribute to learners' resistance to AI tools, which are often perceived as impersonal, autonomous, and lacking human mentorship. Emotional factors such as technostress and lack of self-confidence may also be amplified by an educational culture that values teacher-led instruction and rigorous error correction. Furthermore, contextual concerns like resistance to change could be intensified by limited exposure to advanced educational technologies in some Iranian learning environments. While these findings offer rich insights, transferability to other cultural settings should be approached cautiously. In contexts where AI tools are more normalized, or where student-centered learning is more deeply ingrained, psychological barriers may manifest differently. Therefore, future research could investigate cross-cultural differences in AI adoption to better understand how cultural norms mediate learners' acceptance of AI-driven assessments.

Conclusion

This mixed-methods study investigated the Iranian EFL learners' psychological barriers that hinder the adoption of AI-driven language assessment tools. The findings identified three primary categories of reasons: emotional, perceptual, and contextual factors, each with specific sub-themes. Emotional factors, including lack of self-confidence, and technophobia or technostress, were ranked as the most significant by students. Perceptual factors, such as learners' negative past experiences, lack of trust in AI tools accuracy, and lack of perceived usefulness, followed in importance. Contextual factors, including resistance to change, and privacy issues, were also found to play a role.

The findings of this study suggest several actionable recommendations and implications for different stakeholders to facilitate the adoption of AI-driven language assessment tools. For EFL learners, developing digital literacy and AI-related skills through guided training and hands-on experiences can help reduce technophobia and build confidence in using AI assessments. Educators can support this transition by gradually integrating AI tools into assessment practices, providing clear explanations of their benefits, and fostering a supportive learning environment that addresses students' emotional and perceptual concerns. Encouraging reflective discussions about past experiences with technology and offering personalized feedback can also help mitigate skepticism and negative perceptions.

AI tool developers should focus on designing user-friendly, transparent, and adaptive assessment systems that build trust among learners. Enhancing the accuracy and explainability of AI-generated results can address concerns about reliability, while incorporating learner-centered features can improve perceived usefulness. Policymakers, in turn, should establish clear guidelines for ethical AI implementation, ensuring data privacy and security while promoting professional development programs for educators. By addressing these barriers collaboratively, all stakeholders can contribute to a more effective and inclusive integration of AI in language assessment.

This study, while offering valuable insights, has certain limitations. The relatively small sample size (30 participants) limits the generalizability of the findings to a broader EFL learner population. Additionally, the study focused solely on learners' psychological barriers, overlooking the perspectives of educators, institutions, and other stakeholders who play a crucial role in AI adoption. To build on these findings, future research should include larger and more diverse samples to enhance the reliability and applicability of the results. Furthermore, longitudinal studies could provide deeper insights into how learners' attitudes toward AI-based assessment evolve over time. Experimental studies may also be beneficial in assessing the effectiveness of interventions, such as AI literacy programs or confidence-building strategies, in reducing psychological barriers and fostering greater acceptance of AI-driven language assessment tools.

Statements and Declarations:

Ethical Approval: This study was conducted according to the guidelines of the Declaration of Helsinki and reviewed and approved by the institutional review board (IRB) at Kashmar Institute of Higher Education, Iran. This university does not have a dedicated ethics committee.

Consent to Publish: This research takes informed consent to participate and to publish were obtained from the respondents regarding voluntarily participating in this research.

Data Availability Statement: Not applicable.

Informed Consent: Participants gave written informed consent for review and signature before starting the data collection.

Competing interests: We have no conflicts of interest to disclose.

Funding: No funding organizations

Acknowledgments: Not applicable

References

Abdellatif, M. S., Alshehri, M. A., Alshehri, H. A., Hafez, W. E., Gafar, M. G., & Lamouchi, A. (2024). I am all ears: listening exams with AI and its traces on foreign language learners' mindsets, self-competence, resilience, and listening improvement. *Language Testing in Asia*, *14*(1). https://doi.org/10.1186/s40468-024-00329-6

Ahmad, M., Subih, M., Fawaz, M., Alnuqaidan, H., Abuejheisheh, A., Naqshbandi, V., & Alhalaiqa, F. (2024). Awareness, benefits, threats, attitudes, and satisfaction with AI tools among Asian and African higher education staff and students. *Journal of Applied Learning & Teaching, 7*(1). https://doi.org/10.37074/jalt.2024.7.1.10

Al-khreshehm, M. (2024). The future of artificial intelligence in English language teaching: Pros and cons of ChatGPT implementation through a systematic review. *Language Teaching Research Quarterly, 43*, 54-80. https://doi.org/10.32038/ltrq.2024.43.04

Al-Mughairi, H., & Bhaskar, P. (2024). Exploring the factors affecting the adoption AI techniques in higher education: insights from teachers' perspectives on ChatGPT. *Journal of Research in Innovative Teaching & Learning*. https://doi.org/10.1108/jrit-09-2023-0129

Almuhanna, M. A. (2024). Teachers' perspectives of integrating AI-powered technologies in K-12 education for creating customized learning materials and resources. *Education and Information Technologies*. https://doi.org/10.1007/s10639-024-13257-y

Alqahtani, T., Badreldin, H. A., Alrashed, M., Alshaya, A. I., Alghamdi, S. S., bin Saleh, K., ... & Albekairy, A. M. (2023). The emergent role of artificial intelligence, natural learning processing, and large language models in higher education and research. *Research in Social and Administrative Pharmacy, 19*(8), 1236-1242. https://doi.org/10.1016/j.sapharm.2023.05.016

Biju, N., Abdelrasheed, N. S. G., Bakiyeva, K., Prasad, K. D. V., & Jember, B. (2024). Which one? AI-assisted language assessment or paper format: an exploration of the impacts on foreign language anxiety, learning attitudes, motivation, and writing performance. *Language Testing in Asia*, *14*(1). https://doi.org/10.1186/s40468-024-00322-z

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Cai, Q., Lin, Y., & Yu, Z. (2023). Factors influencing learner attitudes towards CHATGPT-Assisted Language learning in Higher Education. *International Journal of Human-Computer Interaction*, *40*(22), 7112–7126. https://doi.org/10.1080/10447318.2023.2261725

Chan, C. K. Y., & Lee, K. K. W. (2023). The AI generation gap: Are Gen Z students more interested in adopting generative AI such as ChatGPT in teaching and learning than their Gen X and millennial generation teachers? *Smart Learning Environments*, *10*(1). https://doi.org/10.1186/s40561-023-00269-3

Dakakni, D., & Safa, N. (2023). Artificial intelligence in the L2 classroom: Implications and challenges on ethics and equity in higher education: A 21st century Pandora's box. *Computers and Education Artificial Intelligence*, *5*, 100179. https://doi.org/10.1016/j.caeai.2023.100179

Dashtestani, R., & Mohamadi, A. (2023). In-service teachers' conceptualization of implementing interactive learning in online and face-to-face teaching contexts: challenges and opportunities. *Interactive Learning Environments*, 1–12. https://doi.org/10.1080/10494820.2023.2255240

Davis, F. D. (1989). Technology acceptance model: TAM. *Al-Suqri, MN, Al-Aufi, AS: Information Seeking Behavior and Technology Adoption*, *205*(219), 5.

Du, Y. and Gao, H. (2022). Determinants affecting teachers' adoption of AI-based applications in EFL context: An analysis of analytic hierarchy process. *Education and Information Technologies*, *27*(7), 9357-9384. https://doi.org/10.1007/s10639-022-11001-y

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., . . . Williams, M. D. (2019). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, *57*, 101994. https://doi.org/10.1016/j.ijinfomgt.2019.08.002

Guo, S., Shi, L., & Zhai, X. (2025). Developing and validating an instrument for teachers' acceptance of artificial intelligence in education. *Education and Information Technologies*. https://doi.org/10.1007/s10639-025-13338-6

Hazzan-Bishara, A., Kol, O. & Levy, S. (2025). The factors affecting teachers' adoption of AI technologies: A unified model of external and internal determinants. *Education and Information Technologies*. https://doi.org/10.1007/s10639-025-13393-z

Huang, F., Wang, Y., & Zhang, H. (2024). Modelling generative AI acceptance, Perceived Teachers' enthusiasm and Self-Efficacy to English as a Foreign Language learners' Well-Being in the Digital era. *European Journal of Education*. https://doi.org/10.1111/ejed.12770

Kamalov, F., Calonge, D. S., & Gurrib, I. (2023). New era of Artificial intelligence in Education: Towards a sustainable Multifaceted Revolution. *Sustainability*, *15*(16), 12451. https://doi.org/10.3390/su151612451

Kleinheksel, A. J., Rockich-Winston, N., Tawfik, H., & Wyatt, T. R. (2020). Demystifying content analysis. *American Journal of Pharmaceutical Education*, *84*(1), 7113. https://doi.org/10.5688/ajpe7113

Lin, H., & Chen, Q. (2024). Artificial intelligence (AI) -integrated educational applications and college students' creativity and academic emotions: students and teachers' perceptions and attitudes. *BMC Psychology*, *12*(1). https://doi.org/10.1186/s40359-024-01979-0

Lochmiller, C. R. (2021). Conducting thematic analysis with qualitative data. *The Qualitative Report*, *26*(6), 2029-2044. https://doi.org/10.46743/2160-3715/2021.5008

Löhr, K., Weinhardt, M., & Sieber, S. (2020). The "World Café" as a Participatory Method for Collecting Qualitative Data. *International Journal of Qualitative Methods*, *19*. https://doi.org/10.1177/1609406920916976

Maheshwari, G. (2023). Factors influencing students' intention to adopt and use ChatGPT in higher education: A study in the Vietnamese context. *Education and Information Technologies*, *29*(10), 12167–12195. https://doi.org/10.1007/s10639-023-12333-z

Malik, M. A., & Amjad, A. I. (2024). AI vs AI: How effective are Turnitin, ZeroGPT, GPTZero, and Writer AI in detecting text generated by ChatGPT, Perplexity, and Gemini?. *Journal of Applied Learning and Teaching*, *8*(1). https://doi.org/10.37074/jalt.2025.8.1.9

Mogavi, R. H., Deng, C., Kim, J. J., Zhou, P., Kwon, Y. D., Metwally, A. H. S., . . . Hui, P. (2023). ChatGPT in education: A blessing or a curse? A qualitative study exploring early adopters' utilization and perceptions. *Computers in Human Behavior Artificial Humans*, *2*(1), 100027. https://doi.org/10.1016/j.chbah.2023.100027

Owan, V. J., Abang, K. B., Idika, D. O., Etta, E. O., & Bassey, B. A. (2023). Exploring the potential of artificial intelligence tools in educational measurement and assessment.

Eurasia Journal of Mathematics Science and Technology Education, *19*(8), em2307. https://doi.org/10.29333/ejmste/13428

Pillai, R., Sivathanu, B., Metri, B., & Kaushik, N. (2023). Students' adoption of AI-based teacher-bots (T-bots) for learning in higher education. *Information Technology and People*, *37*(1), 328–355. https://doi.org/10.1108/itp-02-2021-0152

Rajak, L., Chauhan, S., & Bara, S. (2024). Transforming English Pedagogy with Artificial Intelligence: Enroute to Enhanced Language Learning. In *Artificial Intelligence: A Multidisciplinary Approach towards Teaching and Learning* (pp. 216-241). Bentham Science Publishers. https://doi.org/10.2174/9789815305180124010013

Rezaei, M., Pironti, M., & Quaglia, R. (2024). AI in knowledge sharing, which ethical challenges are raised in decision-making processes for organisations? *Management Decision*. https://doi.org/10.1108/md-10-2023-2023

Sarte Jr, A. E., & Quinto, E. J. M. (2024). Understanding the importance of weight management: a qualitative exploration of lived individual experiences. *International Journal of Qualitative Studies on Health and Well-being, 19*(1), 2406099. https://doi.org/10.1080/17482631.2024.2406099

Sejnowski, T. J. (2024). *ChatGPT and the Future of AI: The Deep Language Revolution*. Cambridge, Massachusetts: MIT Press.

Selten, F., & Klievink, B. (2023). Organizing public sector AI adoption: Navigating between separation and integration. *Government Information Quarterly*, *41*(1), 101885. https://doi.org/10.1016/j.giq.2023.101885

Stöhr, C., Ou, A. W., & Malmström, H. (2024). Perceptions and usage of AI chatbots among students in higher education across genders, academic levels and fields of study. *Computers and Education Artificial Intelligence*, *7*, 100259. https://doi.org/10.1016/j.caeai.2024.100259

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, *16*(1). https://doi.org/10.1186/s41239-019-0171-0

Zhao, D. (2024). The impact of AI-enhanced natural language processing tools on writing proficiency: an analysis of language precision, content summarization, and creative writing facilitation. *Education and Information Technologies*. https://doi.org/10.1007/s10639-024-13145-5

Zheng, Y., & Stewart, N. (2024). Improving EFL students' cultural awareness: Reframing moral dilemmatic stories with ChatGPT. *Computers and Education Artificial Intelligence*, *6*, 100223. https://doi.org/10.1016/j.caeai.2024.100223

Biodata

Alireza Maleki, a PhD holder in Applied Linguistics, is a certified teacher trainer, researcher, and university lecturer in Khorasan-e Razavi province, Iran. He runs various workshops regarding teaching and assessing English skills interactively in ELT institutions and organizations. He is a journal reviewer and also has a keen interest in educational assessment concepts and subjects. His scholarship has appeared in Journal of Academic Ethics, International Journal of Lifelong Education, Discover Sustainability, The Journal of Asia TEFL, etc.

Appendix 1: Online Focus Group Discussion Questions

The online focus group discussions comprised basic and guiding questions, which included:

1. What are your general thoughts on using AI tools (e.g., automated grading, chatbots) for language assessment?
2. Have you used any AI-based language assessment tools in the past? If so, what was your experience with them?
3. How do you think AI-based language assessment tools compare to traditional assessment methods (e.g., teacher-based, paper exams)?
4. Do you feel comfortable using AI-based language assessment tools in your studies? Why or why not?
5. What are some challenges or barriers you have faced in adopting AI tools for language assessment?
6. What psychological factors influence your perception of AI-based language assessment tools?

Appendix 2: Ranking Scale

Instruction: Based on the themes identified in the online focus group discussions, we kindly ask that you prioritize the following reasons. Please assign a unique rank to each theme by inserting the appropriate number in the provided box, ensuring that each rank is used only once. If you encounter difficulty in deciding between two categories, please use your personal judgment to make the best selection. Please note that all responses will remain confidential and will be used solely for research purposes.

Lack of Trust in AI Tools Accuracy ☐

Technophobia or Technostress ☐

Privacy Issues ☐

Lack of Perceived Usefulness ☐

Lack of Self-confidence ☐

Resistance to Change ☐

Negative Past Experiences ☐