



Investigating Iranian EFL Teacher AI Assessment Literacy: A ChatGPT-Driven Collaborative Reflective Approach

Rajab Esfandiari^{1*}

Mohammad Hossein Arefian²

Abstract

The emergence of artificial intelligence (AI) has prompted teachers to integrate AI tools into assessment practices. Accordingly, teachers need to learn how to implement AI tools to inform their assessment practices with a sufficient level of AI-based assessment literacy. One promising approach to developing teachers' AI-based assessment literacy can be collaborative reflective practice (CRP), particularly when facilitated by tools such as ChatGPT or other generative AI technologies. This study qualitatively explored how English-as-a-foreign-language (EFL) teachers can reflect and collaborate with ChatGPT, individually and collectively, through CRP to develop their AI-based assessment literacy. After collecting qualitative data from six novice and experienced EFL teachers through observations, group discussions, and interviews within a descriptive case research design, we found that EFL teachers developed their AI-based assessment literacy emotionally, socially, and professionally through ChatGPT-driven CRP. Teachers reflected on their AI-based assessment practices, collaborated with ChatGPT to identify and solve problems related to using AI tools in assessment, received feedback and guidance from ChatGPT on using AI tools in their assessment practices, and improved their knowledge of AI-based assessment through ChatGPT-driven CRP. The findings contribute to the growing dialogue on AI integration in professional development by empirically showing that ChatGPT-driven collaborative reflective practice can serve as a viable, accessible, and context-responsive mechanism for developing EFL teachers' AI-based assessment literacy, particularly in under-resourced contexts where institutional professional development opportunities are limited.

Keywords: AI-based assessment literacy, ChatGPT-driven CRP, collaborative reflective practice, professional development

* Review History:

Received: 05/07/2026

Revised: 20/07/2026

Accepted: 19/08/2026

1. Professor, Department of English Language Teaching, Faculty of Literature and Humanities, Imam Khomeini International University, Qazvin, Iran. (Corresponding Author) rbesfandiari@gmail.com
2. Ph.D. Candidate, Department of English Language Teaching, Faculty of Literature and Humanities, Imam Khomeini International University, Qazvin, Iran. arefian.m@yahoo.com

How to cite this article:

Esfandiari, R. & Arefian, M.H. (2026). Investigating Iranian EFL teacher AI assessment literacy: A ChatGPT-Driven collaborative reflective approach. *Teaching English as a Second Language Quarterly*, 45(4), 29-53.
<https://doi.org/10.22099/tesl.2026.55936.3494>



Language assessment practices have changed considerably over the past few decades, requiring updated knowledge and skills about the latest developments in alternative assessments, educational technology, and formative assessment strategies (Morris et al., 2021). English as a Foreign Language (EFL) teachers must have sufficient assessment literacy to use different assessment methods effectively to monitor and improve students' language competence continuously. Despite the popularity of paper-and-pencil assessments, the expansion of innovative technological tools and platforms has encouraged teachers to use artificial intelligence (AI) tools in their language assessment practices. EFL teachers use AI tools, such as automated essay scoring systems, adaptive testing platforms, and data-driven learning analytics, to improve the quality of their assessment tasks, reduce the burden of grading, and enhance their AI-based assessment literacy (Hooda et al., 2022).

Given the significance of AI in assessment practices, EFL teachers need an adequate level of AI-based assessment literacy. This literacy refers to the knowledge, skills, and understanding teachers need to implement AI tools effectively, ethically, and sustainably (Almatrafi et al., 2024). This knowledge is essential for teachers to design reliable, authentic, and valid assessment tasks effectively, administer assessment tasks through AI technology, build e-portfolios to record learners' performance, interpret assessment results automatically, and monitor, measure, and improve students' language learning via AI tools (González-Calatayud et al., 2021). Given AI's growing significance in assessment, EFL teachers must improve their AI-based assessment literacy (Dakakni & Safa, 2023). EFL teachers can develop it by acquiring and deepening the knowledge and skills needed to navigate AI tools proficiently and incorporate AI applications into assessment practices pedagogically, ethically, and technically (Almatrafi et al., 2024). Nevertheless, AI-based assessment practices must follow ethical principles, social justice, human rights, and other humanistic factors to respect teachers' cognitive, emotional, autonomous, and initiative-driven capabilities (Cukurova & Miao, 2024).

EFL teachers can improve their AI-based assessment literacy through professional development approaches, such as collaborative reflective practice (CRP) (Ding et al., 2024). Through this approach, teachers can critically reflect on their AI-based assessment practices and experiences with peers to build their AI-based assessment competence, knowledge, and ability (Celik, 2023). Engaging in CRP can help EFL teachers explore and share their AI-based assessment practices and ideas, provide peer feedback on how they use AI tools to assess, and develop an understanding of how to implement AI tools to improve their assessment practices. They can fine-tune their previous assessment practices with AI tools and make necessary changes to enhance their AI-based practices reflectively and collaboratively in a professional community of practice (Du & Gao, 2022). To gain more insightful input, immediate feedback, and constructive guidance, EFL teachers can do CRP with ChatGPT, "a natural language processing system developed by OpenAI (Deng & Lin, 2022, p.81).

Although these challenges are global, the Iranian EFL context presents different structural barriers to professional development. Restricted access to international platforms due to

intermittent sanctions, a centralized educational system that limits instructional autonomy, and minimal institutional support for formative teacher training generate a context where conventional collaborative professional development is often logistically impossible (Afshar & Doosti, 2022). In such settings, where teachers lack time and supervisory guidance, generative AI tools like ChatGPT can serve a critical function. By offering on-demand, individualized scaffolding and reflective prompts outside institutional constraints, ChatGPT serves as a digital mentor available anytime and anywhere—unlike human supervisors, who are often overburdened or inaccessible. This accessibility makes ChatGPT a distinctively viable mediator for collaborative reflective practice in contexts like Iran, where teachers frequently navigate professional isolation alongside technological and pedagogical challenges. This study seeks to understand how ChatGPT-mediated CRP works within these specific limitations.

Teachers can conduct CRP with ChatGPT to develop their AI-based assessment literacy individually anytime and anywhere, and collaboratively within communities of practice (Sidorkin, 2024). Because ChatGPT can provide timely information, respond to questions and issues promptly, and offer recommendations on AI-based assessment practices, it may serve as an indispensable professional mediator for conducting CRP individually when colleagues are not available within a community of practice (Gill et al., 2024). Collectively, EFL teachers can conduct ChatGPT-driven CRP to improve their AI-based assessment practices within their communities of practice, leading to more systematic and informative collaboration. EFL teachers can use ChatGPT-driven CRP to build their AI-based assessment knowledge and skills, prompt them to notice specific AI-based assessment practices and experiences, and help them identify the benefits and limitations of AI tools in assessment (Bower et al., 2024).

Few studies have explored how teachers' AI-based assessment literacy could be developed through ChatGPT-driven CRP, both individually and collectively in a community of practice. This gap matters because it leaves the dynamic interplay between an extremely accessible generative AI tool and the socially situated process of teacher collaboration unexamined. This combination could democratize professional development in under-resourced settings. Accordingly, this study investigates how EFL teachers used ChatGPT-driven CRP to enhance their AI-based assessment literacy independently and collectively.

Literature Review

EFL Teachers' AI-Based Assessment Literacy

Assessment is a significant aspect of language teaching and learning and is estimated to take up 30% to 50% of teachers' time (Brown & Hudson, 1998). To conduct professional assessment practices, EFL teachers need sufficient assessment literacy. They need to select and develop appropriate assessment methods, conduct assessments efficiently, and score and grade tests fairly and reliably. In addition, teachers should analyze results and communicate them effectively to students, parents, and school managers, make reliable and informed decisions, and recognize unethical assessment practices (Gao et al., 2021). Thus, EFL teachers are

expected to know what to assess, how to assess, and why to assess (Xu & Brown, 2016). They should understand assessment concepts in different assessment methods by developing a sound knowledge base on language assessment (i.e., knowledge). Moreover, they need the competence and ability to construct reliable assessment items, score students' performance accurately, use statistics, and interpret and share results effectively (i.e., Skills). Additionally, teachers must be familiar with the principles and rationale for using particular assessment methods and approaches to enhancing fairness, reliability, and validity (i.e., principles) (Du & Gao, 2022).

Modern and innovative technological tools, such as AI, have affected traditional paper-and-pencil assessment practices by assessing learners' language achievement and ability (Dimitriadou & Lanitis, 2023). AI has revolutionized assessment, shifting it toward a more systematic, objective, engaging, formative, and modern approach, though not without criticism (González-Calatayud et al., 2021). Hence, EFL teachers are expected to develop AI-based assessment literacy, meaning they know how to use AI tools to design, implement, and evaluate assessment tasks (Alharbi, 2023). AI-based assessment literacy refers to the understanding and competence of teachers, students, and policymakers to efficiently conduct and analyze assessments driven by AI tools (Sperling et al., 2024). It involves knowledge of the ethical, technical, and pedagogical implications of using AI in educational assessment. Teachers need to understand how AI works, makes decisions, and generates data to integrate AI tools into assessment practices effectively and to critically evaluate AI-based assessment tools for reliability, appropriateness, and accuracy (Celik, 2023).

They need to use AI tools to develop and select appropriate AI-based assessment methods, understand the principles behind AI algorithms, conduct, score, and interpret results with AI tools, make reliable and valid decisions, and understand unethical AI-driven assessment practices (Celik, 2023; González-Calatayud et al., 2021). Moreover, they can design assessment tasks, provide immediate feedback, record students' performance through portfolios, address assessment challenges, and enhance fairness, authenticity, and meaningfulness by implementing AI tools in their assessment practices (Koraishi, 2023). Thus, EFL teachers must know how to apply AI tools in assessment practices, understand how AI algorithms work, and evaluate and interpret AI-generated information (Ayanwale et al., 2024). When teachers possess a sufficient level of AI-based assessment literacy, they can use AI tools to offer real-time and contextualized feedback, conduct adaptive testing and assessment, score essays automatically, modify assessments to meet students' needs, check pronunciations, have a writing assistant, and perform formative assessment (Huang et al., 2023). To develop formative assessment with AI tools, teachers can offer personalized and immediate feedback via automated essay scoring, assess students based on their needs and abilities through adaptive learning platforms, record students' learning through e-portfolios, assess skills and subskills with AI-powered language learning platforms, and conduct self-assessment to recognize and improve their weaknesses using AI tools.

Still, the usefulness of AI tools fundamentally rests on teachers' literacy in implementing them efficiently. Unfortunately, teachers' AI-based assessment literacy has received relatively limited scholarly attention (Fang et al., 2023). Professional development programs need to equip EFL teachers, novice or experienced alike, with fundamental AI-based assessment literacy by adopting various approaches and methods, including attending webinars and conferences, engaging in joint publications, and participating in reflective and collaborative practices (Jiang, 2022).

ChatGPT-Driven Collaborative Reflective Practice for Professional Development

AI, which can process human-oriented data to acquire, analyze, correct, and use it for language-learning tasks and activities (Popenici & Kerr, 2017), can enhance, change, and develop EFL teachers' assessment and instructional practices. It involves tutoring systems, intelligent agents, and collaborative learning systems, offering numerous benefits for EFL teachers, such as a tailored professional learning environment, real-time feedback, enhanced motivation, instant availability, and meaningful inclusion (Jiang, 2022). More recently, ChatGPT has emerged as an effective AI tool for developing instructional materials, producing human-like audio and text, interpreting linguistic patterns to provide feedback, identifying concepts and theories, solving problems efficiently, and acting as a personalized tutoring system that can improve instruction and assessment (Kasneci et al., 2023). ChatGPT can be seen as a virtual partner that supports EFL teachers by providing timely guidance, knowledge, and skills to help them improve their teaching and assessment practices reflectively and collaboratively (Lisnawati & Muharam, 2023). EFL teachers can also reflect and collaborate through CRP with ChatGPT to improve AI-based assessment literacy by reflecting on experiences, collaborating to receive support and feedback, identifying and solving problems, developing new pedagogical strategies, and building awareness of implementing AI tools in assessment. In this case, they can identify and resolve issues, support their professional development, and build competence in their AI-based assessment practices through ChatGPT-driven CRP (Celik, 2023; Kohnke et al., 2023). Although some investigations have examined ChatGPT's influence on language learning and teaching (Jeon & Lee, 2023), it is crucial to explore how ChatGPT-driven CRP can improve EFL teachers' AI-based assessment knowledge, abilities, and practices.

Through ChatGPT-driven CRP, EFL teachers can reflect on their assessment practices using AI tools to identify AI-based assessment problems and work toward improving them collectively and reflectively. They can plan strategies, implement them, and observe and reflect on their AI-based assessment practices collectively within a community of practice through ChatGPT-driven CRP. They can look for resources for assessment tasks, invite AI to provide feedback on their self-constructed rubrics, ask AI to give feedback on student writing, and compare and contrast AI-generated and human-generated assessment interpretations (Bower et al., 2024). EFL teachers can engage in CRP with AI-driven ChatGPT outside of school to use their free time and develop AI-based assessment literacy more autonomously. They can engage in ChatGPT-driven CRP on their AI-oriented assessment practices with colleagues within a

community of practice at their school to receive emotional, cultural, and social support that individual reflection with ChatGPT does not provide (Liu, 2024). Therefore, this study aims to examine how ChatGPT-driven CRP can foster EFL teachers' AI-based assessment literacy.

Research Gap and Contribution

Prior research, as reviewed in the previous parts, establishes three key strands of knowledge. First, earlier studies show that AI-based assessment literacy includes technical, instructional, and ethical dimensions (González-Calatayud et al., 2021; Almatrafi et al., 2024), and that EFL teachers necessitate structured professional development to acquire this literacy (Ding et al., 2024). Second, collaborative reflective practice has been shown to nurture teachers' assessment literacy by facilitating peer dialogue, critical reflection, and shared problem-solving (Du & Gao, 2022; Şener & Mede, 2023). Third, existing research has documented ChatGPT's potential as a virtual mentor that can offer timely feedback, personalized scaffolding, and on-demand guidance for teachers (Kasneci et al., 2023; Kohnke et al., 2023; Sidorkin, 2024).

The intersection of these three domains remains critically underexplored. Whereas each area has received independent scholarly attention, no empirical study to date has systematically explored how ChatGPT can be intentionally incorporated into collaborative reflective practice to develop teachers' AI-based assessment literacy. More specifically, the literature presents four significant gaps. First, most studies on AI-based assessment literacy have focused on conceptualizing the construct or measuring teachers' current competence, with insufficient attention to how this literacy can be developed through continuous professional development interventions (Fang et al., 2023). Second, while CRP has been broadly studied as a professional development approach, most CRP research remains limited to human-to-human collaboration, leaving the role of AI mediators largely unexplored (Celik, 2023). Third, existing ChatGPT research in education has mainly focused on student learning outcomes or teachers' pedagogical practices, with assessment—despite being a core instructional responsibility—receiving disproportionately little attention (Jeon & Lee, 2023). Lastly, the specific contextual challenges of under-resourced contexts like Iran, where institutional support is restricted and access to international platforms is limited (Afshar & Doosti, 2022), have not been sufficiently addressed in the AI and assessment literacy literature.

This study directly addresses these gaps by exploring how EFL teachers can use ChatGPT-driven CRP—both individually and collectively within a community of practice—to foster their AI-based assessment literacy. Unlike previous research that has examined AI tools and CRP in isolation, this study adopts an integrated approach that positions ChatGPT not simply as a technical tool but as a reflective partner and collaborative mediator. By doing so, the present study contributes to the developing body of knowledge on AI incorporation in teacher professional development in three ways: (1) empirically documenting a process through which ChatGPT-driven CRP fosters AI-based assessment literacy; (2) recognizing the emotional, social, and professional benefits and challenges of this approach; and (3) providing

contextualized insights from the Iranian EFL setting, a context that remains underrepresented in the global AI and language assessment literature.

Theoretical Framework

This study draws on a theoretical framework linking Loughran's (2002) conception of reflective practice and the Technological Pedagogical Content Knowledge (TPACK) framework (Mishra & Koehler, 2006). Loughran's model views reflection as a purposeful, structured process containing not just recollection but rigorous analysis of practice to inform future action. This supports our operationalization of CRP. Concurrently, the TPACK framework conceptualizes the knowledge teachers need to incorporate technology effectively. We posit that 'AI-based assessment literacy' is a specific manifestation of TPACK, where Technological Knowledge (TK) includes understanding AI tools like ChatGPT, Pedagogical Knowledge (PK) includes assessment principles, and Content Knowledge (CK) is English language proficiency. Developing AI-based assessment literacy thus involves the reflective and collaborative negotiation of these knowledge domains, mediated by an AI tool. The following two research questions are used to guide this qualitative study:

1. To what extent can EFL teachers enhance their AI-based assessment literacy through ChatGPT-driven CRP?
2. What are the potential benefits and challenges when the teacher informants conduct ChatGPT-driven CRP to promote their AI-based assessment literacy?

Method

Research Design

This study selected a qualitative approach within a participatory action research (PAR) framework, using a descriptive case study methodology to document the process. Grounded in the principles of PAR (Kemmis et al., 2013), this research involved a structured intervention in which the researcher actively supported the ChatGPT-driven CRP process. This interventionist approach was vital to collaboratively produce new knowledge and practical solutions for developing AI-supported assessment literacy within the participants' particular context. The 'case' in this study is defined as the bounded system of six EFL teachers engaging in this researcher-facilitated professional development intervention over four months. A qualitative, case-oriented design was chosen to produce rich, contextualized data on the complex processes and outcomes of this novel practice.

Participants and Setting

This qualitative case study included three novice and three experienced EFL teachers (three male and three female). With fewer participants, the study could dedicate more time to exploring participants' perspectives, practices, and perceptions of how novice and experienced teachers used ChatGPT-driven CRP to improve their AI-based assessment literacy. Novice and

experienced EFL teachers were also selected to provide diversity, collect diverse insights, enhance transferability, identify common patterns and unique differences, build contextual understanding, and understand challenges and benefits at different career stages (Creswell, 2013). This study explored their perceptions, experiences, procedures, practices, and the benefits and challenges of conducting ChatGPT-driven CRP to enhance their AI-based assessment literacy. The teachers were between 24 and 38 years of age (Mean = 30), and all were Iranian. Following Bennett et al.'s (2013) classification, novice teachers had 1 to 3 years of experience, and experienced teachers had 7 to 15 years of experience.

All graduated in language teaching or literature and earned Bachelor of Arts (BA) degrees from state-run and private universities in Iran. All teachers worked at the same language institute, teaching EFL to adult learners with proficiency levels from A1 to C1 according to the Common European Framework of Reference (CEFR) (Little, 2006). Standard textbooks adopted communicative language teaching approaches published by Cambridge University Press. Because they all worked in the same language institutes as EFL teachers, the lead researcher purposively selected this context, as they used online teaching platforms, prepared online learning materials, conducted online assessments, and used AI tools to improve students' language learning. Accordingly, the context was suitable for this study. This cohort of participants was selected through purposive sampling because they were familiar with AI tools, willing to use AI in their assessment practices, and experienced in conducting online assessments. They also needed to know how to develop professionally through reflective practices. These criteria ensured that participants had the necessary background to engage meaningfully with the ChatGPT-driven CRP process. Although the sample size is small and homogeneous, limiting the generalizability of the findings, it was selected for the depth and richness of data it could offer.

Data collection continued over four months until saturation was reached. Saturation was determined through a continuous, iterative process of data analysis that ran simultaneously with data collection (Braun & Clarke, 2006). After each round of group discussions and interviews, the lead researcher transcribed and preliminarily coded the data and then held weekly debriefing sessions with the research team to review emerging codes and themes. Saturation was deliberately reached when three conditions were met: (a) no new codes appeared from two consecutive group discussions or interviews, (b) present themes were adequately rich and well-supported by data from different participants, and (c) the research team reached consensus that additional data collection was unlikely to yield novel insights relevant to the research questions. Evidence supporting this determination involved a declining rate of new code generation—from an average of 12 new codes per session in the first month to 0–2 new codes per session in the final month and the consistent recurrence of established themes across all six participants' accounts. This systematic approach ensured data collection continued only until informational redundancy was reached, supporting best practices for saturation in qualitative research (Saunders et al., 2018).

Data Collection

To guarantee a comprehensive and triangulated data collection process, three main instruments were used: semi-structured interviews, observations, and recordings of group discussions. An interview protocol was adopted to probe teachers' experiences, perceptions, and the perceived outcomes of the ChatGPT-driven CRP process (see Appendix A). The protocol included open-ended questions such as "Can you describe a particular instance where a reflection with ChatGPT changed your approach to an AI-based assessment?" and "What were the most significant challenges you encountered when collaborating with peers and ChatGPT on assessment tasks?" The protocol was piloted with an EFL teacher who met the study's criteria but did not participate. This pilot led to minor improvements in question clarity and sequencing.

A structured observation rubric was integrated to document teachers' application of AI tools in their assessment practices (see Appendix B). The rubric focused on three key areas: the teacher's technical implementation of AI tools in assessment tasks (for instance, incorporating automated scoring or generating quiz items), their instructional interactions with students during AI-assisted assessments, and any verbalized reflective comments the teacher made about the process. As a participant-observer, the researcher used this rubric to take detailed field notes during each observed session.

The researcher held sixteen weekly group discussions over the four-month intervention period and used a set of directing prompts to initiate reflective dialogues. Prompts involved "Describe a challenge you faced this week integrating an AI tool for assessment and how you and ChatGPT addressed it" and "Share an example of feedback you gained from ChatGPT on your assessment rubric and discuss with the group how you plan to use it." All group discussions were audio-recorded and transcribed verbatim for analysis.

Data Collection Procedures

Participants were chosen purposefully according to the criteria outlined in the Participants section. The lead researcher offered an initial two-hour training session to present AI tools for assessment, such as automated essay scoring and adaptive testing platforms, as well as procedures for implementing CRP with peers and ChatGPT. This session was teaching only and outlined the process without intervening in the baseline data.

Following the training, the four-month intervention began. Participants were instructed to incorporate AI tools into all aspects of their assessment practices, from design and administration to evaluation and interpretation. They were required to engage in ChatGPT-driven CRP in two modes: individually, by reflecting on their practices with ChatGPT outside of school hours, and collectively, by participating in bi-weekly group discussions where they shared insights, exchanged experiences, and collaboratively reflected on their AI-based assessment practices with peers and the researcher. During these group discussions, participants were motivated to integrate ChatGPT in real time to seek feedback, produce ideas, and deepen their reflective dialogue.

The researcher functioned as a participant-observer throughout this process. In the group discussions, the researcher fostered conversation, offered prompts, asked critical questions, and recorded the sessions. Each teacher was formally observed three times, at the beginning, middle, and end of the intervention, for one hour per session. These observations, directed by the structured rubric, offered direct evidence of how teachers' AI-based assessment practices developed over time. We took detailed field notes during these observations.

At the end of the four months, we conducted individual semi-structured interviews with each participant. These interviews, approximately one hour each, were designed to collect in-depth, retrospective accounts of their experiences, the perceived impact on their AI-based assessment literacy, and the benefits and challenges of the ChatGPT-driven CRP process. All interviews were conducted in English and audio-recorded for transcription.

Data Analysis Procedures

Thematic analysis, following Braun and Clarke's (2006) six-step framework, was used to analyze the transcribed data from interviews, group discussions, and observation field notes. The process was iterative and recursive. The lead researcher constantly read all transcripts and notes to become immersed in the data. The team then produced initial codes manually, identifying salient features relevant to the research questions. To establish intercoder reliability and alleviate individual bias, a second coder, an experienced colleague in qualitative research and applied linguistics, individually coded a sample of approximately 20 percent of the data. The coders then discussed their coding until they reached consensus, achieving a 90 percent agreement rate on the final coded sample. This collaborative process improved the codebook, which the lead researcher then applied to the full dataset.

The following phases included grouping codes into potential themes, reviewing these themes against the coded extracts and the full dataset to ensure coherence, and defining and naming each theme. The coding process followed a hybrid approach, combining deductive and inductive elements. Deductive codes were primarily derived from the theoretical framework and research questions, involving codes related to TPACK components (Technological Knowledge, Pedagogical Knowledge, Content Knowledge) and reflective practice dimensions (e.g., 'planning,' 'acting,' 'observing,' 'reflecting'). Inductive codes developed organically from the data, capturing unanticipated patterns including 'emotional responses to ChatGPT,' 'over-reliance on AI,' and 'cultural misalignment of AI recommendations.'

The initial coding phase produced 87 distinct codes across all data sources. Through iterative refinement—collapsing redundant codes, merging conceptually similar codes, and removing codes with insufficient supporting evidence—these were consolidated into 12 sub-themes and subsequently organized into 4 overarching final themes: (1) ChatGPT-driven CRP and AI tools reform assessment practices creatively; (2) ChatGPT-driven CRP helps develop reflective and collective AI-based assessment literacy; (3) emotional benefits and challenges of ChatGPT-driven CRP; and (4) social and professional benefits and challenges of ChatGPT-driven CRP.

These final themes, presented in the Findings section, show the most significant and consistent patterns related to developing AI-based assessment literacy through ChatGPT-driven CRP.

Ensuring Trustworthiness

Different strategies were implemented to guarantee the trustworthiness of the analysis, framed by the criteria of transferability, credibility, dependability, and confirmability (Lincoln & Guba, 1985). To improve credibility, triangulation was used by comparing evidence from three data sources: interviews, observations, and group discussions. Moreover, member checking was implemented by sharing a summary of the primary findings with all participants, who confirmed that the interpretations accurately reflected their experiences. Transferability was addressed through thick, descriptive accounts of the participants, research context, and ChatGPT-driven CRP process, allowing readers to assess the potential applicability of the findings in other contexts.

Dependability was supported by a detailed description of the research design, data collection instruments, and analytical procedures, creating an audit trail that allows others to follow and assess the consistency of the research process. Lastly, confirmability was established by maintaining a reflexivity journal documenting the lead researcher's assumptions, preconceptions, and evolving perspectives, which were consciously bracketed during analysis. The research team also actively searched for and discussed negative cases that contradicted developing themes, ensuring that interpretations remained robust, balanced, and grounded in the data.

Results

The First Research Question 1: To what extent can EFL teachers enhance their AI-based assessment literacy through ChatGPT-driven CRP?

ChatGPT-Driven CRP and AI Tools Reform Assessment Practices Creatively.

AI tools offer new ways to conduct assessments for different purposes: monitoring student learning, improving learning, and measuring language ability by systematizing and innovating assessment practices. For instance, NT1 used ChatGPT to produce vocabulary test items, explaining that

I gave ChatGPT a particular prompt: 'Create 10 multiple-choice vocabulary questions for intermediate EFL learners concentrating on academic words from Unit 3, with plausible distractors.' ChatGPT generated a range of questions, and then we reviewed and adapted them to fit my students' proficiency level and the specific vocabulary they had learned.

Similarly, ET2 applied ChatGPT to create a speaking assessment rubric, stating that

I asked ChatGPT: 'Generate a rubric for assessing EFL speaking skills at the B2 level, focusing on fluency, accuracy, pronunciation, and vocabulary, with four proficiency levels.' ChatGPT produced a detailed rubric with descriptors for each level. I, then,

compared it with my present rubric, adapted the descriptors to my instructional context, and shared it with my colleagues for feedback.

NT2 used ChatGPT to offer feedback on student writing. She commented that

I pasted a student's essay into ChatGPT and asked: 'Offer constructive feedback on this B1-level essay, focusing on grammar, vocabulary, and organization, and recommend three areas for improvement.' ChatGPT provided detailed feedback, which we then reviewed, personalized to the student's needs, and shared with the student.

ET3, faced with the challenge of grading 30 speaking recordings, asked ChatGPT: *"How can we effectively assess and offer feedback on 30 student speaking recordings without spending excessive time?"* ChatGPT recommended using a structured observation checklist with key criteria. ET3 described the process as follows: *"I adapted the checklist ChatGPT proposed and used it to streamline my grading process, saving significant time while preserving consistency across all students."*

These examples show how teachers used ChatGPT to address concrete assessment challenges, produce assessment materials, and receive guidance on improving their assessment practices. As novice teacher 3 (NT3) presented: *"When we used AI tools to enhance, monitor, and improve students' writing ability, we realized how we could use assessment continuously to enhance students' writing skills with innovative tasks, constructive feedback, and helpful self- and peer-assessment throughout the semester."*

AI could prepare, conduct, and evaluate assessment tasks more reliably and objectively, based on established assessment principles and theories. In this respect, Novice teacher 1 argued that

"AI could potentially guide both novice and experienced teachers in assessing their students more professionally and creatively with various assessment tools, approaches, methods, and practices, such as pop quizzes, role-play, language puzzles, flashcard storytelling, and so more. To assess vocabulary, teachers learned how to construct reliable, objective, and meaningful assessment tasks and items, implement them professionally, ethically, and efficiently with online platforms, and evaluate, score, and interpret them to show progress, identify weaknesses, and provide feedback through AI tools."

However, a clear disagreement emerged in how novice and experienced teachers engaged with this creative potential. Novices often assumed AI's suggestions were general, as they needed an experiential framework to evaluate them critically. For example, while NT1 enthusiastically listed AI-produced assessment types, they did not, in following observations, articulate how these activities aligned with particular learning objectives. In contrast, experienced teachers like ET3 worked as critical filters. ET3 reported dismissing ChatGPT's suggestion for a gamified vocabulary assessment, explaining, *"The AI proposed a complex point system that would have taken 20 minutes just to explain. We took the core idea, competitive*

quiz, but simplified the rules based on what we know works for my students.' This demonstrates how experienced teachers leveraged their PK to adapt the TK offered by ChatGPT, a nuanced incorporation not yet evident among novices.

Novice teachers lacked experience with different assessment methods and could not put assessment concepts and theories into practice. Novice teachers, who often struggle to bridge theory and practice, found AI tools particularly useful in linking theoretical principles to practical assessment design. For instance, NT 3 noted that *"AI taught me to use assessment tasks, like peer-assessment, oral task completion, and role-plays to provide the opportunity for authentic language use, feedback, and scaffolding with communicative and task-based language testing"*.

Both novice and experienced EFL teachers could use ChatGPT-driven CRP to improve their AI-based assessment practices by acting, observing, and reflecting on their AI-based assessments collectively in a professional community of practice. Thus, those who used ChatGPT-driven CRP learned how to implement AI tools for various assessment purposes more reflectively and collectively. Novice teacher 1 admitted, *"I learned to reflect collaboratively with ChatGPT in order to use AI to make a grammar test more efficiently by aligning the objectives, students' abilities and needs, and content and context. AI showed me how to make tests, implement them, and interpret the results."* Also, because CRP can enhance teachers' critical and creative thinking skills (Şener & Mede, 2023), novice teachers used ChatGPT-driven CRP critically and creatively, either individually outside school or collectively with colleagues within the school context, to improve their assessment practices, knowledge, and attitudes while implementing AI tools. NT2, for example, mentioned: *"Quizlet, Duolingo, and Kahoot showed me how to assess students practically. Doing CRP with others and ChatGPT in an innovative and critical way helped me learn how to conduct various modes of assessment ethically, effectively, and appropriately."*

ChatGPT-driven CRP served as a model for teaching novice teachers how to construct assessment items and tasks professionally, objectively, and reliably using AI tools. Novice teacher 1 stated, *"When we talked with ChatGPT, it recommended me to use Grammarly to check students' writing mistakes and errors, which helped me save time, and we could rest more after the classes."* Likewise, when novice teachers used ChatGPT reflectively, they acquired and deepened their AI-based assessment literacy and became able to use AI tools for developing, implementing, and interpreting the results of the assessment, as commented by NT3: *"I progressively grasped how to design, conduct, and analyze formative assessment practices by means of AI tools when we used ChatGPT to reflect on my AI-related assessment practices"*. Similarly, they became more confident in generating a personalized assessment by focusing on students' needs, strengths, weaknesses, and areas for improvement after being assisted by ChatGPT. As NT2 highlighted, *"Previously, I was anxious to construct assessment items and score them. However, doing CRP with ChatGPT and other teachers and using AI*

tools helped me to be more assertive and autonomous. Also, we could see students' abilities and change my future teaching to improve them."

ChatGPT-driven CRP also helped experienced EFL teachers change and transform their traditional assessment practices and adopt more innovative assessment practices through AI tools. In this case, experienced teachers used performance-based assessment approaches, such as task-based assessment, alternative assessment, portfolio and dynamic assessment, and learner-oriented assessment, as recommended by ChatGPT. As ET3 observed: *"I acted, monitored, and reflected on my assessment practices with ChatGPT and could use AI to prepare assessment tasks for learning purposes. I could involve students through authentic assessment tasks and provide immediate feedback to enhance students' achievement."* Additionally, experienced teachers could change their assessment practices by using adaptive and alternative assessment methods, providing more constructive and innovative feedback to students, and interpreting results. In the words of ET2,

When we used ChatGPT to reflect, we understood how we could make assessments for learning purposes and improvement through various modes and approaches of AI-based assessments. I developed a series of speaking tasks, engaged students in them, and continuously provided feedback on their pronunciation and intonation, vocabulary and grammar, fluency and accuracy, and discourse and functions.

Experienced EFL teachers could also use ChatGPT-driven CRP to reinvest and co-construct their knowledge of AI-based assessment, gain more knowledge of assessment methods and theories, and modify their assessment practices creatively via AI tools, as pointed out by ET1: *"I could learn new concepts and theories regarding language assessment and testing and learn to develop the fairness and authenticity of my assessment tasks through ChatGPT"*.

ChatGPT-Driven CRP Helps Develop Reflective and Collective AI-Based Assessment Literacy.

ChatGPT-driven CRP enhanced novice and experienced EFL teachers' AI-based assessment literacy through reflective and collaborative practice. A novice teacher stated: *"I learned how to use AI tools in my assessment practice by gaining feedback from ChatGPT. ChatGPT helped me to use AI tools more effectively"* (NT2). Doing ChatGPT-driven CRP with others in communities of practice and individually during independent reflections provided a supportive learning environment to share ideas, receive feedback, construct knowledge, and change practices related to AI-based assessment. As NT2 explained: *"Doing reflections and collaborations with ChatGPT could support teachers to use AI effectively to monitor and gather accurate data to make informed decisions."* ChatGPT could enhance novice teachers' reflective and collaborative practices by asking critical questions, providing helpful solutions, and offering knowledge of recent theories on implementing AI tools in assessment. NT1 stated that *"ChatGPT was a real mentor for me as we could reflect on using AI in my assessment practices."*

And it helped me to solve the problems related to using a new AI tool within assessment practices."

Furthermore, novice teachers gained more benefits when they used ChatGPT with other teachers. Experienced teachers could provide emotional and social support and draw on their established assessment experience in their specific context. NT2 noted, *"When I was sad about my lack of knowledge about making motivating and authentic assessment tasks, an experienced teacher encouraged me to reflect on and try new approaches, which was very supportive."*

Conducting CRP with ChatGPT helped novice and experienced EFL teachers investigate and reflect on their AI-based assessment practices and ideas, share experiences with other experienced teachers, receive feedback, and become more aware of how to use AI tools in their assessment practices. Experienced EFL teachers could reflect on their traditional assessment practices with novice teachers and ChatGPT in a community of practice to add innovation and creativity via AI tools, as mentioned by ET1: *"When we reflected on my reading assessment items with ChatGPT and novice teachers, we learned how to select the appropriate AI tools to generate assessment items and reading passage, implement them, and evaluate them through various AI tools"*. Some experienced teachers were nervous to admit their limited AI literacy. They reflected collaboratively on ChatGPT individually, enabling them to gain knowledge and address their problems using AI tools in their assessment practices. As ET2 argued, *"ChatGPT could help me a lot. It introduced different AI tools, explained how to apply them, and evaluated my AI-based assessment practices."* Despite these reports of support, the data also indicated a significant challenge: a strong reliance on AI that threatened teacher autonomy. This was mainly noticed among novice teachers. NT1's statement, *"ChatGPT was a real mentor for me... it helped me to solve the problems,"* while positive, hints at outsourcing problem-solving. This was made explicit by another novice, NT2, who admitted, *"After a few weeks, we found ourselves unable to draft a simple reading comprehension question without first consulting ChatGPT. we had lost confidence in my own ideas."* This over-reliance signals a negative case within our theme of development, suggesting that without structured support, ChatGPT-driven CRP can unintentionally stifle the teacher creativity and self-efficacy it aims to foster. It underscores the delicate balance between using AI as a scaffold and becoming dependent on it.

The Second Research Question: Benefits and Challenges of Conducting ChatGPT-Driven CRP on AI-Based Assessment Practices

For the systematic, clear presentation of the findings related to the benefits and challenges, the results were categorized into three main types: (a) emotional benefits and challenges, (b) social benefits and challenges, and (c) professional benefits and challenges. Each category is presented below with illustrative quotes and examples from the participants.

Emotional Benefits and Challenges

Doing ChatGPT-driven CRP to enhance their AI-based assessment practices offered numerous emotional benefits for novice and experienced EFL teachers. Emotionally, teachers

became more motivated and confident and reduced their anxiety as they collaborated and reflected with ChatGPT to strengthen their AI-based assessment practices, as expressed by NT1: *"I became motivated and able to continuously use AI tools to monitor, evaluate, and improve students' learning."* Novice EFL teachers received immediate and real-life feedback on their emotions and feelings through ChatGPT whenever they needed an expert or mentor to improve their feelings, leading to lower assessment anxiety. Additionally, ChatGPT could provide emotional support, scaffolding, and mutual guidance for novice teachers to learn about AI-based assessment, as NT2 asserted: *"ChatGPT was the same as a real teacher trainer to me since we could get help, feedback, and knowledge and skill of how to assess with AI tools."*

Similarly, experienced EFL teachers were more resilient and open to experiencing and using ChatGPT to enhance their AI-based assessment practices. Because ChatGPT offered a wide range of perspectives, experienced teachers could reflect on different viewpoints and expand their understanding of AI-based assessments. This could help prevent burnout and attrition among experienced teachers. *"As I was developing continuously through ChatGPT, we learned new things related to AI tools and became interested in applying them in my future assessment practices,"* reflected ET2. Moreover, ChatGPT could offer teachers a personalized learning atmosphere to use AI tools effectively in their assessment practices, as ET1 stated: *"ChatGPT helped me for various purposes. I used it to personalize my learning and reflect on my assessment practices individually."* However, ChatGPT sometimes failed to work effectively and provide relevant responses, which could make teachers frustrated and stressed. Some teachers feared AI would replace them, leading to reluctance to use it. *"We became stressed out when ChatGPT could not understand our point and intention during interaction and provided an irrelevant response (ET1)."*

These frustrations were not just technical but often instructional and ethical, directly engaging the TPACK framework. When ChatGPT offered irrelevant responses, it often failed to incorporate the particular CK of the English language or the PK of the local setting. The most noticeable example came from ET2, who was 'shocked' by a recommendation to implement 'rap songs' for speaking assessment. This was not just a cultural misfire but a failure of the AI to align its TK with the cultural and instructional realities of the Iranian classroom. This negative case is essential because it indicates that teachers' development of AI-based assessment literacy is not just about learning to use the tool, but also about developing a critical faculty to identify and reject its misapplications, a key element of reflective practice.

Social Benefits and Challenges

Socially, novice and experienced EFL teachers could build a professional community of practice to reflect collectively, build social bonds, share experiences, link theory and practice, exchange ideas, and make sense of their AI-based assessment experiences through ChatGPT. As ET1 emphasized: *"We could develop together in the community by interacting with ChatGPT to talk about AI tools, solve related problems, and give feedback in order to improve our AI-based assessments."* In addition, ChatGPT-driven CRP could enhance their inclusion

and engagement within communities of practice by giving them opportunities to develop their assessment literacy professionally. ET2 shared the following: *"We were more engaged with developing and enhancing our assessment practices through AI tools."* However, some teachers could not provide the exact prompts and hints when interacting with ChatGPT, and it wrongly evaluated or interpreted social events, phenomena, and meanings. Likewise, teachers relied heavily on AI and ChatGPT, which could weaken their skills in managing social communications, problem-solving, and independence, as maintained by NT1: *"I could not use my own innovation and creativity to produce answers, solve problems, develop lesson plans, and implement and evaluate actions concerning AI-based assessment without the help of ChatGPT."* This statement from a novice teacher contrasts with experienced teachers' reports of integrating ChatGPT as a sounding board or a source of raw ideas to develop further, further highlighting how career stage mediates the risk of over-reliance.

Professional Benefits and Challenges

Professionally, novice and experienced EFL teachers reflected collaboratively and individually with ChatGPT to enhance their critical professional skills and competence regarding their AI-based assessment practices. They increased their digital literacy, assessment literacy, and AI literacy by identifying and solving problems, taking new actions, enhancing their awareness, and building their knowledge and experience through ChatGPT. NT2, for example, remarked that *"ChatGPT could help me find ways to enhance my knowledge and practice of AI-based assessment practices through reflection, collaboration, and action."* Because ChatGPT prompted deeper, more systematic reflective practice, novice and experienced teachers could make more informed decisions and improve their AI-based assessment practices. However, ChatGPT recommended some practices or strategies that were unethical and unsuitable for the context of Iran, as highlighted by ET2: *"When ChatGPT told me to use AI tools to assess speaking through rap songs and rhythms, I was shocked as this cannot be the case for public schools in Iran."*

Discussion

In this study, EFL teachers learned to acquire, deepen, and develop their AI-based assessment literacy by conducting ChatGPT-driven CRP in a community of practice. They used ChatGPT to reflect collectively, build social bonds, share experiences and ideas, and make sense of their AI-based assessment practices. AI tools helped teachers understand, design, conduct, and evaluate assessments in innovative ways. This finding is in line with that of González-Calatayud et al. (2021), who found that AI tools could empower teachers to design customized assessments and offer detailed feedback. However, our study extends their work by showing that ChatGPT, more specifically, functions not only as a technical tool but also as a reflective partner that improves collaborative and individual reflection on assessment practices. Unlike previous studies which primarily focused on the technical capabilities of AI tools

(Hooda et al., 2022), the findings of this study highlight the instructional and reflective dimensions of AI incorporation in assessment. This may be because of the capability of AI tools to analyze extensive amounts of data and offer meaningful insights into learning patterns, strengths, and weaknesses of students (González-Calatayud et al., 2021).

The findings indicate that through ChatGPT-driven CRP, EFL teachers learned to implement AI tools for more personalized assessment practices. For example, one participant used ChatGPT to generate customized vocabulary test items by identifying her students' proficiency level and the specific content they had studied. Likewise, another reported using ChatGPT to provide individualized writing feedback, which she then adapted to each student's needs. An experienced teacher used ChatGPT's recommendations to design an efficient observation checklist for assessing 30 speaking recordings, allowing him to provide consistent yet individualized evaluation. These teacher-reported practices suggest that ChatGPT-driven CRP helped teachers move beyond generic assessment approaches toward more tailored, student-responsive practices that addressed learners' particular strengths, weaknesses, and learning contexts.

Because AI tools serve as mentors that help teachers develop assessment literacy, EFL teachers received guidance and direction from AI tools to design tests and tasks, implement assessments, analyze and interpret assessments accurately and reliably, and share results professionally. Because AI-based assessment literacy was a prerequisite for integrating AI tools into assessment practices, EFL teachers enjoyed using ChatGPT-driven CRP to reflect and collaborate with others, and ChatGPT to enhance their AI-based assessment literacy (Sidorkin, 2024; Wang et al., 2023).

These findings can be interpreted through the incorporated lens of our theoretical framework. ChatGPT-driven CRP operationalized Loughran's (2002) model of reflective practice, guiding teachers beyond simple recollection toward critical analysis of their AI-based assessment actions. Moreover, the development of "AI-based assessment literacy" demonstrates the dynamic negotiation of the TPACK framework (Mishra & Koehler, 2006). Teachers were not just acquiring TK about AI tools but actively integrating it with their PK of assessment principles and CK of the English language. This incorporation, mediated by reflective dialogue with an AI partner, underlines the complex knowledge synthesis needed for effective AI integration in education, as Celik (2023) notes. This finding aligns with Celik's (2023) conceptualization of Intelligent-TPACK, which postulates that teachers need to develop knowledge of how AI tools interact with instructional and content knowledge. Still, the present study empirically showed that this knowledge is not developed in isolation but through a dynamic, socially situated process of collaborative reflection. While Celik's work was mainly theoretical, our findings provide empirical evidence of how TPACK negotiation happens in practice when teachers engage in ChatGPT-driven CRP with peers. Moreover, our results extend the TPACK framework by showing that emotional and social dimensions—for example, reduced anxiety, increased confidence, and peer support—play a crucial role in this knowledge

negotiation, aspects that have received limited attention in previous TPACK research (Mishra & Koehler, 2006).

The EFL teachers did ChatGPT-driven CRP to recognize, prepare, and use AI tools in their assessment practices for various purposes. They learned which AI tools to use and how to use the tools effectively before implementing them in real assessment practices by interacting with ChatGPT. Also, reflecting on and collaborating with ChatGPT individually at home and collectively with colleagues helped EFL teachers develop their AI-based assessment literacy. They were able to link assessment theories and practices, contextualize assessment practices based on students' needs and abilities, and enhance their knowledge of assessment principles such as reliability, validity, authenticity, and fairness by using AI tools. Moreover, novice EFL teachers learned to make their assessment practices more creative, systematic, and objective by interacting with and reflecting on ChatGPT collectively. They acquired the ability to plan and use AI tools effectively, implement them strategically, reflect on the use of AI for assessment consciously, identify and solve problems through interacting with ChatGPT, and take further actions. These findings echo those of Kasneci et al. (2023) and Al-khresheh (2024), who also found that ChatGPT could positively foster teachers' personalized and dynamic learning interactions, produce educational content, nurture student engagement and collaboration, and personalize their learning experiences. However, our study extends these findings by showing that ChatGPT's benefits go beyond individual teacher learning to include collective professional development within a community of practice. While Kasneci et al.'s (2023) study was focused mainly on the opportunities and challenges of large language models for education in general, our study specifically explored how EFL teachers used ChatGPT for assessment purposes, a domain that has received limited attention in previous research (Jeon & Lee, 2023). Additionally, unlike Al-khresheh (2024), who examined teachers' perspectives on incorporating ChatGPT in English language teaching from a global lens, our study offers contextualized insights from the Iranian EFL context, where institutional support and access to international platforms are limited (Afshar & Doosti, 2022).

When EFL teachers designed, administered, and analyzed assessments with AI tools, they reflected on the challenges, collaborated with others and ChatGPT to receive feedback, transformed their practices, enhanced their AI and assessment literacy, and became more critical through CRP with ChatGPT. In this way, they developed their strengths and mitigated weaknesses in their AI-based assessment practices through ChatGPT-driven CRP. Experienced EFL teachers who had established strong assessment knowledge, experience, and skill could modify their assessment knowledge, practices, and beliefs by using AI tools and doing ChatGPT-driven CRP, leading to more adaptability and flexibility to use AI tools in their assessment practices. Moreover, experienced EFL teachers also have more resources or social capital to draw on when choosing what AI tools to use and how to use them to support their assessment activities. AI tools helped them explore new assessment approaches, use more authentic assessment tasks, focus on constructive feedback, and deepen their knowledge of

formative and alternative assessment practices. Besides, doing CRP individually and collaboratively with ChatGPT provided the foundation to develop their ability to select a suitable AI tool, enhance their AI literacy, conduct and evaluate AI-based assessment practices, and use AI to fulfill multiple educational purposes (Almatrafi et al., 2024; Dakakni & Safa, 2023; Gill et al., 2024). They not only increased their assessment literacy but also developed their AI literacy by recognizing and addressing problems, using novel strategies and actions, and deepening their understanding of AI-based assessments through ChatGPT.

EFL teachers could co-develop their AI-based assessment literacy through conducting CRP with ChatGPT. They learned how to produce reliable and valid assessment items with AI tools, score and grade students' performance accurately, strengthen assessment principles via AI tools, and analyze and share results. ChatGPT-driven CRP helped EFL teachers gain real-time scaffolding from ChatGPT, understand the complexities of working with AI, and strengthen their AI-based assessment literacy in a supportive learning environment. In addition, they could make informed decisions about using AI tools in their assessment practices, ask critical questions, share ideas about AI-based assessment, and gain theoretical and practical knowledge of AI-based assessment by incorporating ChatGPT-driven CRP.

This finding aligns with Sidorkin (2024), who argued that chatbots can function as personalized learning companions in higher education, and with Liu (2024), who emphasized the emotional and social support AI can offer teachers. Our study, however, extends earlier research by empirically showing that the benefits of ChatGPT-driven CRP are not uniform across all teachers; rather, career stage mediates the experience. For example, novice teachers, who lacked enough assessment experience, tended to rely more deeply on ChatGPT for concrete solutions, sometimes leading to over-reliance and diminished autonomy. By contrast, experienced teachers used ChatGPT as a platform, critically evaluating and adapting AI-produced recommendations based on their established instructional knowledge. This finding contradicts Jiang et al.'s (2022) suggestion that experienced teachers may be less open to technological innovation; our results indicate that experienced teachers, while cautious, were more selective and strategic in their use of ChatGPT, leveraging their existing assessment knowledge to filter and adapt AI-produced content. This nuanced finding highlights the need for differentiated professional development approaches that address the distinct needs of novice and experienced teachers.

Experienced EFL teachers could build confidence by using ChatGPT-driven CRP individually to strengthen their implementation of AI tools in assessment, which could help them maintain their reputation while interacting with novice teachers (Jiang et al., 2022). Experienced EFL teachers could also enhance their creativity and innovation in assessment practices by using AI tools and learning to interact with and reflect on ChatGPT (Koraishi, 2023). Experienced EFL teachers became more motivated to remain in the profession, with increased resilience and a positive mindset, as they creatively applied AI tools in their assessment practices and developed their AI-based assessment literacy reflectively and

collaboratively through ChatGPT (Bower et al., 2024). Moreover, novice EFL teachers became more autonomous, motivated, and confident in improving their assessment literacy, taking new assessment actions, and improving their assessment. They gained theoretical and practical assessment experiences by implementing AI tools in their assessment practices and collaborating with ChatGPT to receive novel feedback, resources, and content. They also reduced their anxiety by increasing their knowledge of AI assessment tools and ChatGPT. Additionally, both novice and experienced EFL teachers could plan, act, observe, and reflect collectively in a community of practice by using ChatGPT to foster their AI-based assessment knowledge, principles, and attitudes. In this case, EFL teachers could fill their knowledge gaps, examine their AI-based assessment practices more systematically, and solve their AI-oriented assessment problems innovatively through ChatGPT-driven CRP.

Whereas the findings are promising, this study has significant limitations that must be acknowledged. An important limitation stems from the PAR design and the researcher's dual role as facilitator and observer. This interventionist stance, while essential to initiate the process, may have introduced power dynamics and social desirability bias, in which participants offered responses they believed the researcher wanted to hear. While we used member checking and reflexivity journals to alleviate this, it remains a limitation. Moreover, the study's dependence on self-reported data from a small, purposive sample within a particular Iranian context means the findings demonstrate a potential and a process, rather than offering generalizable outcomes.

The practical challenges of fostering AI-based assessment literacy through ChatGPT-driven CRP, as our data show, include irrelevant, unethical, and unsuitable responses from ChatGPT, overreliance on AI tools, and digital anxiety (Wang et al., 2023). Our findings underscore the basic challenge of aligning a generic AI tool with particular cultural and contextual norms, as seen when ChatGPT recommended assessment strategies unsuitable for the Iranian context. EFL teachers need to be motivated to use their cognition and creativity in assessment practices, rather than relying on AI tools and getting help from ChatGPT (Al-khresheh, 2024). Professional development programs should run workshops on using AI tools in assessment to strengthen EFL teachers' ethical, consequential, and professional practices when assessing learners. Working with ChatGPT can also be another area for EFL teachers to develop their knowledge and skills. Novice and experienced EFL teachers should be trained to provide sufficient, necessary, helpful, and instructive information when interacting with ChatGPT to get the right answer, meet students' needs and desires, and respect contextual expectations.

However, teachers may lack sufficient digital literacy, lack access to digital tools, and view implementing AI tools as a tall order. Moreover, conducting CRP individually and collectively can be a time-consuming, longitudinal process, making it challenging for Iranian teachers (Afshar & Doosti, 2022). Furthermore, ChatGPT may provide feedback based on general data patterns and include biases, since it cannot meet specific contextual, cultural, and social values

and norms. It is imperative to consider issues related to transparency, fairness, and the over-reliance on automated tools that may unintentionally diminish professional autonomy.

Conclusion

The results indicated that when novice and experienced EFL teachers conducted ChatGPT-driven CRP, they could design reliable and authentic assessment items and tasks, implement and assess students using AI tools more effectively, and analyze and interpret the results more competently using AI tools. Based on the results presented and discussed in the previous two sections, we conclude that more professional development programs are needed to enhance novice and experienced EFL teachers' knowledge of ChatGPT-driven CRP and improve their AI-based assessment literacy.

Because this study included a small sample, future studies should use larger samples to improve generalizability. This is a perceptual, descriptive study that primarily relies on self-reported data. Other studies could use quantitative correlations and cause-and-effect designs to measure influence and relationships with other variables. Additionally, more qualitative studies can explore teachers' perceptions of ChatGPT-driven CRP and AI-based assessment literacy across different cultures and educational contexts.

This study's findings have theoretical and practical implications. Theoretically, the study contributes to the growing body of knowledge on AI incorporation in teacher professional development by empirically showing that ChatGPT-driven collaborative reflective practice can serve as an accessible, context-responsive mechanism for nurturing AI-based assessment literacy, especially in under-resourced contexts. Essentially, the results recommend that teacher educators and professional development providers should intentionally integrate generative AI tools like ChatGPT into training programs as on-demand reflective partners that supplement traditional mentorship. The differentiated experiences of novice and experienced teachers underscore the need for career-stage-specific professional development, where novice teachers receive additional scaffolding to prevent overreliance on AI, while experienced teachers are encouraged to evaluate AI-produced recommendations using their established instructional knowledge critically. Furthermore, the challenges recognized—including cultural misalignment of AI responses and digital anxiety—highlight the significance of developing teachers' critical AI literacy, empowering them to use AI tools responsibly, ethically, and contextually. Overall, these implications advocate a balanced, human-centered approach in which technology enhances rather than replaces teacher expertise and reflective practice.

Acknowledgments

We would like to thank the editorial team of TESL Quarterly for granting us the opportunity to submit and publish the current synthesis. We would also like to express our appreciation to the anonymous reviewers for their careful, detailed reading of our manuscript and their many

insightful comments and suggestions. We also acknowledge all the participants who took part in this study.

Declaration of conflicting interests

The authors declare no potential conflicts of interest concerning the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for this article's research, authorship, and/or publication.

References

- Afshar, H. S., & Doosti, M. (2022). Implementing and evaluating a peer-coached EFL teacher professional development program. *Evaluation and Program Planning*, 92, 102096. <https://doi.org/10.1016/j.evalprogplan.2022.102096>
- Al-Khresheh, M. H. (2024). Bridging technology and pedagogy from a global lens: Teachers' perspectives on integrating ChatGPT in English language teaching. *Computers and Education: Artificial Intelligence*, 6, 100218. <https://doi.org/10.1016/j.caeai.2024.100218>
- Alharbi, W. (2023). AI in the foreign language classroom: A pedagogical overview of automated writing assistance tools. *Education Research International*, 2023(1), 4253331. <https://doi.org/10.1155/2023/4253331>
- Almatrafi, O., Johri, A., & Lee, H. (2024). A Systematic Review of AI Literacy Conceptualization, Constructs, and Implementation and Assessment Efforts (2019-2023). *Computers and Education Open*, 100173. <https://doi.org/10.1016/j.caeo.2024.100173>
- Bennett, S. V., Brown Jr., J. J., Kirby-Smith, A., & Severson, B. (2013). Influences of the heart: Novice and experienced teachers remaining in the field. *Teacher Development*, 17(4), 562–576. <https://doi.org/10.1080/13664530.2013.849613>
- Bower, M., Torrington, J., Lai, J. W., Petocz, P., & Alfano, M. (2024). How should we change teaching and assessment in response to increasingly powerful generative Artificial Intelligence? Outcomes of the ChatGPT teacher survey. *Education and Information Technologies*, 1–37. <https://doi.org/10.1007/s10639-023-12405-0>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.
- 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://doi.org/10.1016/j.chb.2022.107468>
- Brown, J. D., & Hudson, T. (1998). The alternatives in language assessment. *TESOL Quarterly*, 32(4), 653–675. <https://doi.org/10.2307/3587999>
- 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>
- Deng, J., & Lin, Y. (2022). The benefits and challenges of ChatGPT: An overview. *Frontiers in Computing and Intelligent Systems*, 2(2), 81–83. <https://doi.org/10.54097/fcis.v2i2.4465>
- Dimitriadou, E., & Lanitis, A. (2023). A critical evaluation, challenges, and future perspectives of using artificial intelligence and emerging technologies in smart classrooms. *Smart Learning Environments*, 10(1), 12. <https://doi.org/10.1186/s40561-023-00231-3>

- Ding, A. C. E., Shi, L., Yang, H., & Choi, I. (2024). Enhancing teacher AI literacy and integration through different types of cases in teacher professional development. *Computers and Education Open*, 6, 100178. <https://doi.org/10.1016/j.caeo.2024.100178>
- Du, Y., & 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>
- Fang, Y., Roscoe, R. D., & McNamara, D. S. (2023). Artificial intelligence-based assessment in education. In *Handbook of artificial intelligence in education* (pp. 485- 504). Edward Elgar Publishing. <https://doi.org/10.4337/9781800375413.00033>
- Gao, R., Liu, J., & Yin, B. (2021). An expanded ethical decision-making model to resolve ethical dilemmas in assessment. *Studies in Educational Evaluation*, 68, 100978. <https://doi.org/10.1016/j.stueduc.2021.100978>
- Gill, S. S., Xu, M., Patros, P., Wu, H., Kaur, R., Kaur, K., ... & Buyya, R. (2024). Transformative effects of ChatGPT on modern education: Emerging Era of AI Chatbots. *Internet of Things and Cyber-Physical Systems*, 4, 19–23. <https://doi.org/10.1016/j.iotcps.2023.06.002>
- González-Calatayud, V., Prendes-Espinosa, P., & Roig-Vila, R. (2021). Artificial intelligence for student assessment: A systematic review. *Applied sciences*, 11(12), 5467. <https://doi.org/10.3390/app11125467>
- Hooda, M., Rana, C., Dahiya, O., Rizwan, A., & Hossain, M. S. (2022). Artificial intelligence for assessment and feedback to enhance student success in higher education. *Mathematical Problems in Engineering*, 2022(1), 5215722. <https://doi.org/10.1155/2022/5215722>
- Huang, X., Zou, D., Cheng, G., Chen, X., & Xie, H. (2023). Trends, research issues, and applications of artificial intelligence in language education. *Educational Technology & Society*, 26(1), 112–131.
- Jeon, J., & Lee, S. (2023). Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. *Education and Information Technologies*, 28(12), 15873–15892. <https://doi.org/10.1007/s10639-023-11834-1>
- Jiang, R. (2022). How does artificial intelligence empower EFL teaching and learning nowadays? A review on artificial intelligence in the EFL context. *Frontiers in Psychology*, 13, 1049401. <https://doi.org/10.3389/fpsyg.2022.1049401>
- Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., ... & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Kemmis, S., McTaggart, R., & Nixon, R. (2013). *The action research planner: Doing critical participatory action research*. Springer Science & Business Media.
- Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *Relc Journal*, 54(2), 537–550. <https://doi.org/10.1177/00336882231162868>
- Koraishi, O. (2023). Teaching English in the age of AI: Embracing ChatGPT to optimize EFL materials and assessment. *Language Education and Technology Journal*, 3(1).
- Lisnawati, S. D., & Muharam, S. (2023). Teacher Professionalism Development Strategy through ChatGPT Support in the Context of Education Management. *Journal of Contemporary Administration and Management (ADMAN)*, 1(3), 150–155.
- Little, D. (2006). The Common European Framework of Reference for Languages: Content, purpose, origin, reception and impact. *Language Teaching*, 39(3), 167–190. <https://doi.org/10.1017/S0261444806003557>
- Liu, J. (2024). ChatGPT: perspectives from human–computer interaction and psychology. *Frontiers in Artificial Intelligence*, 7, 1418869. <https://doi.org/10.3389/frai.2024.1418869>
- Loughran, J. J. (2002). Effective reflective practice: In search of meaning in learning about teaching. *Journal of Teacher Education*, 53(1), 33–43. <https://doi.org/10.1177/0022487102053001004>
- Cukurova, M., & Miao, F. (2024). *AI competency framework for teachers*. UNESCO Publishing.

- Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, 108(6), 1017–1054.
<https://doi.org/10.1111/j.1467-9620.2006.00684.x>
- Morris, R., Perry, T., & Wardle, L. (2021). Formative assessment and feedback for learning in higher education: A systematic review. *Review of Education*, 9(3), e3292.
<https://doi.org/10.1002/rev3.3292>
- Popenici, S. A., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research and practice in technology enhanced learning*, 12(1), 22.
<https://doi.org/10.1186/s41039-017-0062-8>
- Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J., Bartlam, B., ... & Jinks, C. (2018). Saturation in qualitative research: exploring its conceptualization and operationalization. *Quality & Quantity*, 52(4), 1893–1907. <https://doi.org/10.1007/s11135-017-0574-8>
- Şener, B., & Mede, E. (2023). Promoting learner autonomy and improving reflective thinking skills through reflective practice and collaborative learning. *Innovation in Language Learning and Teaching*, 17(2), 364–379. <https://doi.org/10.1080/17501229.2022.2047694>
- Sidorkin, A. M. (2024). *Embracing chatbots in higher education: the use of artificial intelligence in teaching, administration, and scholarship*. Taylor & Francis.
- Sperling, K., Stenberg, C. J., McGrath, C., Åkerfeldt, A., Heintz, F., & Stenliden, L. (2024). In search of artificial intelligence (AI) literacy in Teacher Education: A scoping review. *Computers and Education Open*, 100169. <https://doi.org/10.1016/j.caeo.2024.100169>
- Xu, Y., & Brown, G. T. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, 58, 149–162.
<https://doi.org/10.1016/j.tate.2016.05.010>
- Wang, Y., Pan, Y., Yan, M., Su, Z., & Luan, T. H. (2023). A survey on ChatGPT: AI-generated content, challenges, and solutions. *IEEE Open Journal of the Computer Society*.