Chapter 9

EFL School Teacher's AI Assessment Literacy: An Exploratory Study

Hao Xu

National Research Centre for Foreign Language Education, Beijing Foreign Studies University, Beijing, China

This study investigates the AI assessment literacy of high school EFL teachers using the Chinese AI-powered chatbot, ERNIE Bot. Interviews and digital protocols revealed a dynamic relationship between teachers' practical and conceptual knowledge, with both areas mutually reinforcing each other. However, technical knowledge emerged as a potentially independent domain, operating separately from practical and conceptual understandings. This finding underscores the importance of recognizing technical proficiency as a critical component of AI assessment literacy, alongside the interconnectedness of practical and conceptual knowledge.

Introduction

In the rapidly evolving landscape of applied linguistics, the integration of Artificial Intelligence (AI) technology into language teaching and learning has emerged as a significant and long-lasting trend. This technological revolution has not only transformed the way language is taught and acquired but has also raised critical questions about how language teachers adapt and develop their knowledge and literacy to engage effectively with AI tools. Quite recently, some researchers have begun to explore AI literacy or literacies. For instance, Gupta et al.'s (2024) study examined how discussing metaphors for AI can promote a nuanced understanding of AI systems, particularly large language models like ChatGPT, and contribute to the development of critical AI literacy by analyzing metaphors using a collaborative autoethnographic approach and

Cite as: Xu., H. (2024). EFL school teacher's AI assessment literacy: An exploratory study. In C. A. Chapelle, G. H. Beckett, & J. Ranalli (Eds.), *Exploring artificial intelligence in applied linguistics* (pp. 137-150). Iowa State University Digital Press. https://doi.org/10.31274/isudp.2024.154.09

exploring their implications for anthropomorphism, ethics, equity, accessibility, and the different types of literacies. The findings of this study, which emphasize the importance of metaphor reflection in fostering a nuanced understanding of AI, are particularly relevant to teachers. The intersection between teacher cognition and behavior, as manifested in their engagement with AI technology, is a pivotal aspect of understanding how teachers negotiate the demands of this technological shift. Against this backdrop, this exploratory study aims to investigate the current state of AI literacy among English as a Foreign Language (EFL) teachers, exploring how they perceive, understand, and utilize AI tools in their teaching practice. By examining the intersection of teacher cognition and behavior in the context of AI use, the study seeks to gain insights into the factors that influence the development of AI literacy among EFL teachers and the implications this has for teaching effectiveness and student engagement.

Within the broader context of AI's influence on language teaching and learning, this study specifically focuses on examining how a group of high school EFL teachers utilized the ERNIE Bot to assist them in assessing students' English writing and providing feedback. ERNIE Bot, an advanced AI system, is capable of analyzing language use, offering suggestions for improvement, and even assisting in content creation. By exploring this particular application of AI technology in the EFL classroom, the study aims to gain a nuanced understanding of how teachers engage with and incorporate AI tools into their teaching practices. This study not only aligns with the larger problem of understanding how language teachers develop their AI literacy but also contributes to addressing the practical challenges and opportunities associated with the integration of AI in education. By investigating the use of ERNIE Bot in writing assessment and feedback, this study aims to delve into teachers' assessment literacy when utilizing AI tools, exploring how they integrate AI-driven insights into their evaluation practices and understanding the implications for teaching effectiveness and student learning outcomes.

In this study, AI assessment literacy refers to the ability of EFL teachers to understand, evaluate, and effectively utilize AI tools for the assessment of students' language skills, particularly writing. This literacy encompasses not only the technical skills required to operate AI systems but also the pedagogical knowledge needed to integrate AI-driven insights into teaching practices in a meaningful and educationally beneficial way. Teachers' AI assessment literacy is centered around three pillars of knowledge: conceptual, practical, and technical, which can be summarized as conceptuality, practicality, and technicality. Conceptuality, or conceptual knowledge, refers to teachers' deep understanding of assessment principles and theories. This includes comprehension of the purpose, processes, and criteria involved in effective assessment, as well as an awareness of how these elements intersect with the use of AI tools (Pu & H. Xu, 2023). Conceptuality enables teachers to make informed decisions about assessment design and implementation, ensuring alignment with educational goals and student needs. Practicality, on the other hand, pertains to teacher' hands-on experience and practical wisdom gained through assessment implementation. It involves understanding what works well and what doesn't in real-world classroom settings, based on observations, reflections, and iterative improvements (Pu & H. Xu, 2021). Practicality allows teachers to adapt and refine their assessment practices to meet

the unique needs of their students and the demands of the curriculum. Lastly, technicality encompasses teachers' proficiency in using AI tools for assessment purposes. This includes familiarity with the features and functions of various AI platforms, as well as the ability to navigate their interfaces, customize settings, and troubleshoot any issues that may arise (H. Xu, 2022). Technicality empowers teachers to harness the full potential of AI in enhancing the efficiency and effectiveness of their assessment practices. In summary, teachers' AI assessment literacy is underpinned by a blend of conceptuality, practicality, and technicality. These three elements work together to equip teachers with the knowledge and skills needed to integrate AI into their assessment strategies, thereby enhancing the quality and impact of student assessment. This trifecta of conceptuality, practicality, and technicality serves as the conceptual framework for the study, providing a comprehensive understanding of the knowledge and skills required for teachers to integrate AI into their assessment strategies effectively. The research questions this study addresses are:

1. What practical, conceptual, and technical knowledge do high school EFL teachers show as they use ERNIE Bot for assessment and feedback?
2. How do the three kinds of knowledge interact?

As can be seen, the first question aims to explore the different types of knowledge that high school EFL teachers demonstrate when utilizing ERNIE Bot, a specific AI tool, for assessment and feedback purposes. It seeks to understand the teachers' practical know-how in using the tool, their conceptual understanding of the underlying principles and theories of assessment, and their technical proficiency in operating it. The second question focuses on understanding how the three types of knowledge mentioned in the first question (practical, conceptual, and technical) interact with each other. It explores the relationships and intersections between these knowledge domains and how they jointly contribute to the teachers' ability to effectively use ERNIE Bot for assessment and feedback. The aim is to gain insights into how these different types of knowledge complement and reinforce each other in the context of AI-assisted language assessment.

Methods

This study employed a qualitative multiple-case design to investigate teachers' AI literacy across three dimensions: practicality, conceptuality, and technicality, as well as the interaction between these dimensions. The multiple-case design is particularly suited to this study as it enables a comparison and contrast of teachers' AI assessment literacy within and across cases (Yin, 2018). This approach facilitates the identification of commonalities and differences in how EFL teachers engage with AI, both practically and conceptually, as well as their level of technical proficiency. Furthermore, exploring the interaction between these dimensions provides insights into how they are interconnected and influence teachers' overall AI assessment literacy.

Participants and Context

In selecting participants for this study, special consideration was given to teachers who had previously expressed interest and involvement in related research. Specifically, six EFL teachers

from six distinct high schools in the urban areas of Beijing, China, were chosen based on their participation as questionnaire respondents in a prior study conducted by Pu and H. Xu (2021, 2023). This previous research project focused on teachers' online assessment practices during the COVID-19 pandemic, making these teachers particularly suitable for this follow-up investigation exploring a related aspect of educational technology use. Their familiarity with the topic, combined with their voluntary commitment to participate in further studies, ensured a rich and informed data set for the current study.

All of the six participants had experience using the ERNIE Bot for assessment and feedback purposes, primarily in the context of teaching writing. As trial users, they had access to a version of the bot that included watermarks on their screens to record their identity in the form of a unique sequence of numbers and letters. This allowed for the tracking of their interactions with the bot and ensured that data collected from their usage was properly attributed. Table 1 below shows the details of the six participants.

Table 1

Participants' Details

Pseudonym	Gender	Age	Years of teaching	Degree
Liu	Female	39	17	B.A.
Wang	Female	43	13	M.Ed.
Chong	Male	32	8	M.A.
Su	Female	32	7	M.A.
Yang	Female	33	9	M.Ed.
Li	Female	44	21	B.A.

In terms of their teaching practices, these participants regularly assigned writing tasks to their students. Typically, towards the end of each unit in the textbook, there was a writing task that was closely related to the theme and topic of the unit. With six units per semester, this meant that students were given a major writing task approximately every other week. This provided a consistent and structured context for the teachers to implement the ERNIE Bot and gather feedback on their students' writing skills.

Prior to commencing the study, ethical approval was obtained from the author's institution. Additionally, all of the participants provided formal consent to participate in this study, ensuring that their participation was voluntary and that they were fully aware of the nature and purpose of the research.

Data Collection

The present study employed two primary sources of data to investigate teachers' AI assessment literacy across the three dimensions: practicality, conceptuality, and technicality. These sources were individual interviews and digital protocols. The interviews served as the primary means of

collecting qualitative data. They were specifically tailored to focus on teachers' experiences with using the ERNIE Bot for assessing senior high school students' English writing tasks as outlined in the units of the textbook. Interview questions were carefully crafted to elicit detailed responses about teachers' perceptions of the bot's practical utility, their understanding of the underlying concepts related to AI assessment in the context of English writing instruction, and their technical proficiency in using the tool. To ensure a comprehensive understanding, the researcher conducted three interviews with each participant. The initial interview served as an exploratory session to gather preliminary insights into the teachers' experiences and perceptions. Subsequently, a second/third interview was conducted to follow up on initial findings, allowing for a more in-depth exploration of emerging themes and issues.

The interview guide was carefully designed to facilitate a comprehensive understanding of the teachers' experiences with the ERNIE Bot in assessing students' English writing tasks. The questions were crafted to delve into various aspects of the integration process, including the practicalities of using the bot, its perceived benefits and challenges, and the impact on student attitudes and performance. The guide began with questions aimed at elucidating the teachers' specific methods of incorporating the bot into their assessment routines, requesting examples and detailing the helpfulness of the bot's feedback. This was followed by inquiries into any obstacles encountered and strategies employed to overcome them, providing a nuanced picture of the implementation process. To gain a comparative perspective, the teachers were asked to contrast the bot's assessments with traditional methods, highlighting potential advantages or disadvantages. Changes in student dynamics, such as attitudes and performance, were also explored, aiming to capture any observable shifts resulting from the introduction of the bot. The guide then turned to the teachers' personal insights, querying which aspects of the bot's functionality they found most valuable in the context of senior high school writing assessment. This was supplemented by questions about how the bot had influenced their understanding of AI's broader educational potential, particularly in language assessment. Technical proficiency and confidence in troubleshooting were also addressed, aiming to assess the teachers' comfort level with the technology. Additionally, the guide included questions about the teachers' interactions with their peers, exploring whether and how they had shared their experiences with the bot and the reactions they had received. Finally, the interview concluded with a forward-looking question, inviting suggestions and ideas for improving future versions of the bot to better align with the teachers' assessment needs. This comprehensive approach ensured that the interview guide not only captured the current state of the teachers' experiences but also provided valuable feedback for future development. Example interview questions are shown below:

- How have you been using the ERNIE Bot to assess your students' English writing tasks?
- Can you provide specific examples of how the bot's feedback has been helpful in improving students' writing?
- What challenges have you faced when integrating the bot into your assessment process, and how have you addressed them?

- How do you compare the bot's assessments with traditional methods of assessing writing skills?
- Have you noticed any changes in students' attitudes or performance after introducing the bot into your classroom?
- What aspects of the bot's functionality do you find most beneficial for assessing writing at the senior high school level?
- How has using the bot influenced your understanding of AI's potential in education, particularly in the area of language assessment?
- Do you feel confident in your ability to troubleshoot technical issues related to the bot's use?
- Have you shared your experiences with other teachers, and if so, what has been their reaction?
- What improvements would you like to see in future versions of the bot to better support your assessment needs?

In addition to the interviews, the participants were asked to provide digital protocols in the form of Word files containing their ERNIE bot use logs. These logs documented their interactions with the bot, including the types of writing tasks assessed, the frequency of use, and any noted issues or successes. Altogether, the researcher received 89 documents of these logs from the participants, providing a rich source of data to complement the insights gained from the interviews.

The combination of interview data and digital protocols allowed for a comprehensive understanding of teachers' AI assessment literacy across all of the three dimensions. The interviews provided rich qualitative data on teachers' experiences and perceptions in the context of assessing English writing skills, while the digital protocols offered objective usage data to support these findings. Together, these data sources enabled a nuanced analysis of how teachers engage with and benefit from AI-based assessment tools in their teaching practices.

Data Analysis

The data analysis process comprised three iterative rounds, as outlined by Miles et al. (2020), each designed to explore the three dimensions of assessment literacy: practicality, conceptuality, and technicality.

In the first round, the analysis focused on identifying events and situations relevant to teachers' engagement with the AI assessment tool. Within this context, attention was paid to understanding teachers' practical know-how, conceptual understanding, and technical proficiency. Events were examined to reveal how teachers applied AI assessments in their daily practices, reflecting on their understanding of the tool's practical utility and conceptual foundations, as well as their technical skills in using it.

The second round delved deeper into the knowledge embedded within these events, unpacking teachers' perceptions, challenges, and successes. This involved examining not only

their practical experiences but also their conceptual grasp of the principles underpinning AI assessment and their technical adeptness in navigating the tool's features and functionalities. Through this analysis, themes and patterns emerged that shed light on teachers' assessment literacy across all three dimensions.

Finally, the third round employed a comparative lens to contrast varying perspectives and experiences across participants. This allowed for a broader understanding of how practical, conceptual, and technical aspects of assessment literacy influenced teachers' engagement with the AI tool. By juxtaposing diverse accounts, a richer and more nuanced picture of teachers' assessment literacy emerged, encompassing all three dimensions. A limitation of the study should be noted and acknowledged, i.e., a second coder was absent in the data analysis process, which could have potentially enhanced the trustworthiness of the findings.

Findings and Discussion

In this section, findings are presented in response to the research questions, focusing on teachers' practical, conceptual, and technical knowledge as manifested in their assessment and feedback practices. The analysis reveals how these three kinds of knowledge interacted, painting a comprehensive picture of teachers' assessment literacy. This section highlights key findings that emerged from the data analysis, notably the reciprocal relationship between practical and conceptual knowledge and the prominent features of technical knowledge as a relatively independent, yet integral, knowledge area within the broader landscape of assessment literacy.

Practical and Conceptual Knowledge: Reciprocal Relationship

A key finding emerging from the data analysis is the reciprocal manner in which practical and conceptual knowledge appeared to interact. Teachers' day-to-day assessment practices were not only shaped by their theoretical understanding of assessment but also informed and refined this understanding in turn (Y. Xu & Brown, 2016). This dynamic interplay was evident in the way teachers described their approaches to assessment, feedback, and the challenges they faced. In what follows, I present interview extracts that illustrate this reciprocal relationship, providing insights into how teachers' practical experiences informed their conceptual understanding and vice versa. For instance, Chong demonstrated his practical knowledge in the following extracts:

> I have not only asked the bot to help me identify language errors in students' works, but also asked it to help me reformulate their writings. (Chong, interview 1)

> When their writing is full of errors, error detection will be what I ask it to do; when their writing is weak in logic and coherence, reformulation will be what I need it to do for me. (Chong, interview 2)

The extracts provided by Chong demonstrate a clear understanding of practical knowledge in the context of language assessment and revision. In the first extract, Chong explained that he not only used the bot to detect errors in students' writing but also to help him

reformulate their work. This indicates that Chong had a practical grasp of how to apply technology in order to enhance the quality of students' writing, recognizing that different issues required different approaches. In the second extract, Chong further illustrated his practical knowledge by specifying the types of problems he encountered in students' writing and the corresponding solutions he expected the bot to provide. He mentioned that when writing was full of errors, error detection was the priority, whereas when logic and coherence were lacking, reformulation became necessary. This ability to diagnose specific problems and select appropriate tools or strategies to address them is a hallmark of practical knowledge in action. Overall, Chong's interview extracts reveal a teacher who is not only familiar with the theoretical concepts of language assessment and revision but also able to translate this knowledge into practical applications that directly impact student learning. His awareness of the need for both error detection and reformulation, depending on the nature of the student's writing, demonstrates a sophisticated understanding of the practical demands of teaching and assessing writing skills.

It is important to note that Chong's practical knowledge was underpinned by his conceptual knowledge as can be seen from the following extract:

> Reformulation promotes student's attention. When they read the reformulated text, they need to pay closer attention to its differences from their own text. That's how attention is directed to the stuff [i.e., logic and coherence] that is less easy to notice. (Chong, interview 2)

Chong's practical knowledge, as demonstrated in his use of reformulation to enhance student writing, was firmly underpinned by his conceptual understanding of the learning process. In the extract provided, he explained that reformulation promoted students' attention by asking them to compare their original text with the revised version. This process directed their focus to subtleties that might otherwise be overlooked, such as improvements in logic and coherence. Chong's awareness that reformulation could be a powerful tool for drawing students' attention to these aspects of writing reflected a deep understanding of how learners process and internalize new information. His approach aligned with educational theories that emphasize the importance of active engagement and metacognition in learning, suggesting that he viewed the revision process not merely as a corrective exercise but as an opportunity for students to reflect on and improve their own writing skills. In summary, Chong's practical application of reformulation techniques was rooted in his conceptual knowledge of how students learn and develop as writers. His ability to link theory with practice in this way was a testament to the integrated nature of his professional knowledge and skills.

Similarly, Li's conceptual knowledge and practical knowledge were also closely connected. The following extract shows her conceptual knowledge:

> I think assessment via AI should be best conducted in line with the same guidelines followed in conventional assessment … The purpose of assessment should be clear enough or it needs to be clarified until it is fully understood by the tool … The results

> should be communicated effectively with the students – this should obviously be done by the teacher. (Li, interview 1)

Li's conceptual knowledge, as expressed in the extract, demonstrated a clear understanding of the principles underlying effective assessment practices. Li emphasized the importance of aligning AI-based assessment with conventional assessment guidelines, ensuring that the purpose of the assessment was clear and understood by the tool being used. This suggests that Li had a solid grasp of the fundamental concepts and purposes of assessment in education. Furthermore, Li's statement about the need for effective communication of assessment results to students, ideally through the teacher, reflected a conceptual understanding of the educational process and the role of the teacher as a mediator between students and assessment tools. As can be seen, Li's conceptual knowledge, as evidenced in the extract, revealed a deep understanding of assessment principles and practices that was likely to inform and guide her practical application in an educational setting. The close connection between Li's conceptual and practical knowledge was evident in her ability to articulate clear guidelines for the implementation of AI-based assessment while also recognizing the importance of teacher involvement in communicating results effectively to students.

The following data presentation demonstrates how Li's conceptual knowledge has been translated into practical application:

> I often specify the purpose for the bot. I also use a couple of texts to try to see if it fully understands the purpose, before I ask it to provide feedback for more texts. If it understands well enough, it will give feedback that is relevant to the purpose; if not, its feedback will focus on other issues. (Li, interview 2)

In this extract, we see a clear example of how practical knowledge is derived from conceptual understanding. Li's conceptual knowledge about the importance of specifying a purpose for an AI bot and testing its understanding through trial texts informed her practical approach to using the bot effectively. By first establishing a clear purpose and then assessing the bot's ability to understand and respond to that purpose through a series of tests, Li demonstrated the application of conceptual knowledge to a real-world situation. This practical approach allowed Li to ensure that the bot was providing relevant feedback aligned with the intended purpose. In summary, Li's practical knowledge in using the AI bot was firmly rooted in her conceptual understanding of how to establish a clear purpose and assess the bot's ability to understand and respond to that purpose effectively. This example highlights the interconnectedness of conceptual and practical knowledge and how one informs the other in real-world applications.

Technical Knowledge: Distinct Yet Integrated

Transitioning from our discussion on the interplay between practical and conceptual knowledge, we now turn our attention to the distinct yet integrated realm of technical knowledge:

> I put all of the students' written assignments in one word file, and uploaded it to the tool … I typed in my instruction in English, not in Chinese, as I feel it works better as the instruction is given in the same language as that of the data. (Liu, interview 2)

> I once let the bot read a student's work, asking it to point out the weaknesses of her writing. I knew that there was a lack of grammatical complexity, as she had used very simple sentence structures, even for complex ideas. The bot did not point this out … I then gave the bot a new task, asking it to pretend to be the student who had written the work. It produced very simple sentence structures … The underlying rubrics are not always stated! (Yang, interview 3)

In the extracts provided by Liu and Yang, we see manifestations of technical knowledge in their respective approaches to using AI assessment tools. Liu demonstrated technical proficiency in preparing and uploading students' written assignments to the tool, ensuring a smooth and efficient process. Yang, on the other hand, showcased a deeper understanding of the tool's limitations and how to work around them by repurposing the bot's functions.

These examples highlight how technical knowledge in teachers' AI assessment literacy can be an independent area of expertise. While Liu's technical skills were evident in the effective use of the tool's basic features, Yang's technical knowledge extended beyond mere operation to include an understanding of the tool's underlying mechanisms and limitations. This allowed Yang to adapt the tool to specific assessment needs, illustrating the value of technical knowledge in teachers' ability to customize and enhance AI assessment practices. Thus, the extracts from Liu and Yang not only illustrated how technical knowledge was manifested in the practical use of AI assessment tools but also emphasized its importance as an independent area of knowledge in teachers' AI assessment literacy.

Discussion

So far, we have examined how teachers' practical and conceptual knowledge of assessment and feedback interact in a reciprocal manner, much like in traditional assessment practices (Y. Xu & Brown, 2016). However, it appears that their exposure to AI technologies has not led to the addition of new components to their practical and conceptual knowledge. Instead, technical knowledge has emerged as a distinct body of knowledge that can operate independently, with less attachment to practical and conceptual knowledge. This raises the possibility that technical knowledge might not necessarily promote the development of practical and conceptual knowledge, indicating a need for further exploration of the relationship between these different types of knowledge in the context of AI assessment.

In the EFL teachers' AI writing assessment practices, practical knowledge and conceptual knowledge are seen to be reciprocal, much like in traditional assessment methods. The teachers draw on their conceptual understanding of assessment principles and pedagogy to inform their practical actions, such as providing feedback and evaluating student performance. Conversely, their practical experiences with the AI tool, such as observing its strengths and limitations in

assessing student writing, feed back into their conceptual understanding, refining their beliefs and theories about effective assessment. However, this relationship is uniquely presented as mediated by the AI tool. The AI's automated feedback and evaluation capabilities introduce new practical considerations, such as interpreting the tool's output and determining its alignment with assessment goals. These practical aspects, in turn, require teachers to conceptually reconcile the AI's suggestions with their own understanding of effective writing assessment. This reciprocal relationship between practical and conceptual knowledge, as mediated by the AI tool, highlights the dynamic and iterative nature of teachers' knowledge development in the context of AI-supported assessment.

The possibility that technical knowledge might not necessarily promote the development of practical and conceptual knowledge raises important questions about the relationship between these different types of knowledge in the context of AI assessment. Firstly, the apparent independence of technical knowledge suggests that interaction with AI technology may go beyond merely using it as a tool. If technical knowledge were subordinate to practical and conceptual knowledge, we would expect it to develop in tandem with these other types of knowledge. However, the observed independence of technical knowledge indicates that teachers may need to engage with AI technology in more meaningful ways, beyond just using it for its intended purpose, to fully harness its potential for enhancing assessment practices. Secondly, the training of technical knowledge may need to be reconsidered as an independent field of teacher education. Traditionally, teacher education has focused on developing practical and conceptual knowledge (Y. Xu & Brown, 2016), with technical knowledge often seen as a subordinate skill to be acquired alongside these other types of knowledge (Falloon, 2020). However, in the context of AI assessment, technical knowledge may need to be treated as a distinct area of expertise that requires dedicated training and development. This could involve providing teachers with opportunities to explore the affordances and limitations of AI tools, as well as fostering a deeper understanding of how these tools can be integrated into assessment practices in ways that align with pedagogical goals.

By considering these two points, we can begin to explore the complex relationship between technical, practical, and conceptual knowledge in the context of AI assessment. This exploration is crucial for understanding how teachers can effectively leverage AI tools to enhance their assessment practices and, ultimately, improve student learning outcomes.

Conclusion

In conclusion, this study has delved into the practical, conceptual, and technical knowledge exhibited by high school EFL teachers as they utilize ERNIE Bot for assessment and feedback in the context of L2 writing. We have observed that while teachers possess varying degrees of these three types of knowledge, their effective integration is crucial for optimal use of the AI tool. Practical knowledge, demonstrated through teachers' ability to apply ERNIE Bot's feedback in real-world classroom settings, is enhanced when combined with a solid conceptual understanding of assessment principles and techniques. Technical knowledge, perceived as independent, is essential for troubleshooting and customizing the tool to meet specific learning objectives. The

interplay between these three knowledge domains is evident in teachers' ability to adapt ERNIE Bot's suggestions to individual student needs and to integrate its feedback into their overall assessment strategy. Ultimately, this study underscores the importance of a holistic approach to teacher education that fosters the development of all three knowledge areas, enabling EFL teachers to fully harness the potential of AI-assisted assessment for improved learner outcomes.

The examination of how teachers' practical and conceptual knowledge of assessment and feedback interact in traditional and AI-supported contexts reveals profound implications for teacher education. In particular, the emergence of technical knowledge as a distinct body of expertise, with less attachment to practical and conceptual knowledge, challenges the existing paradigms of teacher training and professional development. Firstly, the finding that practical and conceptual knowledge operate in a reciprocal manner, much like in traditional assessment practices, underscores the importance of fostering this dynamic interaction in teacher education. Teachers need opportunities to develop their conceptual understanding of assessment principles and pedagogy through practical experiences, such as providing feedback and evaluating student performance. At the same time, they must be able to reflect on their practical experiences and refine their beliefs and theories about effective assessment. Teacher education programs should, therefore, incorporate experiential learning activities that allow teachers to engage with assessment practices in authentic classroom settings and debrief with peers and mentors to make sense of their experiences.

Secondly, the observed independence of technical knowledge raises the need for a rethinking of how teachers are trained to engage with AI-supported assessment tools. Teacher education programs must recognize that technical knowledge is not merely a subordinate skill to be acquired alongside practical and conceptual knowledge but a distinct area of expertise that requires dedicated training and development. This training should go beyond basic proficiency in using AI tools and include opportunities for teachers to explore the affordances and limitations of these tools, understand their underlying algorithms and data sources, and critically evaluate their alignment with pedagogical goals. Moreover, teacher education programs should foster a culture of collaboration and co-learning between teachers and AI developers. Such collaboration can help teachers gain a deeper understanding of the technical aspects of AI-supported assessment tools and provide valuable feedback to developers on how these tools can better meet the needs of teachers and students. This, in turn, can lead to the development of more user-friendly and pedagogically sound AI-supported assessment tools.

Finally, the finding that technical knowledge might not necessarily promote the development of practical and conceptual knowledge highlights the need for a holistic approach to teacher education that integrates these different types of knowledge. Teacher education programs should aim to cultivate teachers who are not only proficient in using AI-supported assessment tools but also have a deep understanding of assessment principles and pedagogy and are able to critically reflect on their practices. Such an approach can empower teachers to leverage the potential of AI-supported assessment tools to enhance their practices and improve student learning outcomes.

In closing, our examination of high school EFL teachers' use of ERNIE Bot for assessment and feedback has led us to re-conceptualize AI assessment literacy as an integrated blend of practicality, conceptuality, and technicality. This alternative understanding challenges the prevalent notion that AI assessment literacy is merely the extension of traditional assessment literacy into AI-powered digital environments. The findings underscore the indispensability of technical knowledge, which is often overlooked as a mere adjunct to practical and conceptual understandings. Instead, I argue that technicality co-exists with the AI-powered digitality, forming a permeating culture that shapes teachers' assessment and feedback practices. As AI continues to transform educational landscapes, it is imperative to reconsider teacher education frameworks that nurture a balanced development of practical, conceptual, and technical know-how, enabling teachers to navigate the complexities of AI-assisted assessment with confidence and proficiency.

Acknowledgements

The author thanks the editors and reviewers for their valuable comments. This study was also supported by "Research on the Innovation of Foreign Language Education in China" (grant number: 22JJD740011), a major project of Key Research Institutes of Humanities and Social Sciences under the Ministry of Education, China.

References

Falloon, G. (2020). From digital literacy to digital competence: The teacher digital competency (TDC) framework. *Educational Technology Research and Development, 68*(5), 2449-2472. https://doi.org/10.1007/s11423-020-09767-4

Gupta, A., Atef, Y., Mills, A., & Bali, M. (2024). Assistant, parrot, or colonizing loudspeaker? ChatGPT metaphors for developing critical AI literacies. *Open Praxis, 16*(1), 37-53. https://doi.org/10.55982/openpraxis.16.1.631

Miles, M. B., Huberman, A. M., & Saldaña, J. (2020). *Qualitative data analysis: A methods sourcebook (4th ed.)*. Thousand Oaks, CA: Sage.

Pu, S., & Xu, H. (2021). Examining changing assessment practices in online teaching: A multiple-case study of EFL school teachers in China. *The Asia-Pacific Education Researcher, 30*(6), 553-561. https://doi.org/10.1007/s40299-021-00605-6

Pu, S., & Xu, H. (2023). Equity and teachers' online classroom assessment literacy in China: Insights from parental involvement during the Covid-19 pandemic. *Asia-Pacific Journal of Teacher Education, 51*(4), 381-394. https://doi.org/10.1080/1359866X.2023.2228217

Xu, H. (2022). Interactive self-reporting as behavioural cue elicitor for online-classroom assessment. *Behavioral Sciences, 12*(7), 232. https://doi.org/10.3390/bs12070232

Xu, Y., & Brown, G. T. L. (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

Yin, R. K. (2018). *Case study research and applications: Design and methods (6th ed.)*. Thousand Oaks: Sage.

About the Author

Dr. Hao Xu is professor of Applied Linguistics at National Research Centre for Foreign Language Education, Beijing Foreign Studies University. His main research interests include second language acquisition, language teacher education, and language policy and planning. He is currently associate editor of *International Journal of Applied Linguistics*.