

Bridging Assessment Knowledge and Practice: Pre-Service EFL Teachers' Assessment Literacy in Evaluating Pragmatic Competence

*Entika Fani Prastikawati*¹

*Senowarsito*²

*Dodi Mulyadi*³

*Maria Yosephin Widarti Lestari*⁴

¹ Corresponding author, English Education Department, Universitas Persatuan Guru Republik Indonesia Semarang, Indonesia; entikafani@upgris.ac.id

² English Education Department, Universitas Persatuan Guru Republik Indonesia Semarang, Indonesia

³ English Education Department, Universitas Muhammadiyah Semarang, Indonesia

⁴ English Education Department, Universitas Persatuan Guru Republik Indonesia Semarang, Indonesia

Received: 24 April 2026

Accepted: 27 June 2026

Published: 30 July 2026

Abstract

This study emerges from the recognition that language assessment literacy (LAL) is critical for designing assessments congruent with a communicative language teaching approach to EFL contexts, especially concerning pragmatic competence, an underrepresented area of classroom assessment. Based on the multidimensional LAL framework proposed by Glenn Fulcher, this study investigated pre-service EFL teachers' learning about language (LAL) during teaching practicum. Using an explanatory sequential mixed-methods design, 68 sixth semester pre-service teachers were involved. Overall, LAL was moderate ($M = 2.89$), with moderate scores in both task design ($M = 2.81$) and rubric development ability as well as confidence when creating these tasks associated with high perceived challenges to their implementation ($M = 3.78$). These qualitative results show that much pragmatic competence is evaluated indirectly, focusing more on vocabulary and grammar than appropriateness of content. These results reflect a gap between theory and practice. Therefore, the study points out that more professional-oriented and field-sensitive training is required to reinforce pragmatic assessment in EFL teacher education.

Keywords: Assessment literacy, assessment knowledge, assessment practice, pragmatic competence



To cite this article: Prastikawati, E. F., Senowarsito, Mulyadi, D., & Lestari, M. Y. W. (2026). Bridging assessment knowledge and practice: Pre-service EFL teachers' assessment literacy in evaluating pragmatic competence. *SALEE: Study of Applied Linguistics and English Education*, 7(2), 147–167. <https://doi.org/10.35961/salee.v7i2.2864>

DOI: 10.35961/salee.v7i2.2864

1. Introduction

Language assessment literacy (LAL) is widely acknowledged as an essential aspect of teacher professionalism, especially in English as a Foreign Language (EFL), where assessment procedures significantly influence instruction and learning results. In communicative language teaching, teachers are required to evaluate not only learners' grammatical precision but also their capacity to utilize language suitably inside context (Anam & Putri, 2021; Brown et al., 2024). In numerous EFL classrooms, assessment frequently emphasizes discrete linguistic knowledge, often undermining communicative skills. This disparity illustrates a continual conflict between the objectives of communicative pedagogy and the actualities of classroom evaluation. As noted in recent literature, assessment techniques significantly influence teaching and learning, rendering LAL a crucial domain for teacher development (Fulmer et al., 2015; Tóth & Csapó, 2022). As a result, the capacity of teachers to develop and execute assessments that are meaningful becomes essential for the attainment of more comprehensive educational objectives. The necessity to enhance instructors' assessment proficiency is broadly recognized. Nevertheless, the degree to which this proficiency corresponds to intricate language usage issues remains uncertain.

One area that has received comparatively limited attention in assessment practices is pragmatic competence, which refers to the ability to use language appropriately according to context, social norms, and communicative intention. Researchers such as Afshari et al. (2024) and Hammouri & Al-Khanji (2023) point out, pragmatic competence is one of the aspects of communicative ability, since it helps learners to do speech acts, maintain politeness and interpret what has not been said. Despite its importance, pragmatic competence has received little attention in EFL assessment due to the high level of difficulty it brings when being assessed (Bouknify, 2025; Fathi et al., 2025). Pragmatic performance is context dependent and cannot be scored using fixed procedures like grammar or vocabulary measures because it requires evaluative judgment. Consequently, teachers may either bypass assessing pragmatics or do so using small-scale and informal measures. This tendency generates a gap between the communication competencies learners are supposed to attain and what is actually evaluated in classroom environments. Therefore, cultivating the ability to evaluate pragmatic competence is not simply an ancillary talent but an essential enhancement of instructors' assessment literacy. Resolving this issue necessitates a thorough analysis of teachers' comprehension and implementation of pragmatic assessment in practice.

In this sense LAL is typically viewed as a multidimensional concept that involves some understanding of assessment principles, the ability to design and effectively implement summative and formative assessment-based learning opportunities, alongside capabilities for interpreting assessments and using them appropriately. Giraldo (2021) asserts that effective assessment literacy necessitates the integration of theoretical comprehension with contextual

classroom decision-making. This viewpoint emphasizes that understanding evaluation is inadequate; teachers must also possess the ability to implement this information in contextually relevant manners. Nonetheless, current research suggests that teacher preparation programs frequently prioritize generic evaluation principles while insufficiently covering specific areas like pragmatics (Azizi & Namaziandost, 2023; Esterhazy et al., 2023). Consequently, pre-service teachers may cultivate a superficial comprehension of assessment, while missing the competencies necessary to navigate intricate communicative frameworks. This constraint is most apparent when they must create assignments, formulate rubrics, and assess performance in actual classroom contexts. Thus, the connection between LAL and pragmatic assessment is yet inadequately investigated. Additional inquiry is required to comprehend the practical interactions of these elements.

Although language assessment literacy (LAL) has been conceptualized in various ways, contemporary frameworks consistently portray it as a multidimensional construct encompassing assessment knowledge, practical assessment skills, and contextual or affective dispositions (Fulcher, 2012; Giraldo, 2021; Vogt et al., 2024). Fulcher (2012) emphasizes that language teachers require not only theoretical knowledge of assessment principles but also the ability to design valid assessment tasks, develop scoring criteria, interpret assessment results, and make informed assessment decisions within specific instructional contexts. Likewise, Giraldo (2021) argues that effective LAL extends beyond conceptual understanding to include teachers' confidence in implementing assessments and their ability to respond to contextual challenges encountered during classroom assessment. Drawing on these perspectives, the present study operationalizes LAL into six dimensions relevant to assessing pragmatic competence: (1) knowledge of pragmatic competence, (2) knowledge of assessment principles, (3) pragmatic assessment task design, (4) rubric development skills, (5) confidence in assessing pragmatic competence, and (6) perceived challenges in assessment implementation

The teaching practicum offers a significant framework for analyzing this topic, since it marks the phase when pre-service teachers start to implement their theoretical knowledge in genuine classroom settings. During the practice situations, they are planning lessons, implementing instruction and assessing student learning under the tutelage of mentor teachers (Mankki et al., 2025). These experiences often expose the gap between what they learn in the classroom and what can actually be done. Institutional expectations, time restrictions, and mentorship procedures may affect the execution of assessments. In certain instances, these factors facilitate reflective development, whereas in others, they bolster traditional assessment practices. Consequently, the practicum serves as a pivotal venue for examining the implementation and evolution of LAL. It is also where challenges related to assessing pragmatic competence are likely to become most visible. Understanding this process is essential for identifying areas where teacher preparation may fall short. Therefore, focusing on practicum contexts allows for a more grounded analysis of assessment literacy in action.

Most existing studies on LAL have primarily examined teachers' general assessment knowledge or assessment practices related to discrete language skills, such as reading, writing, or grammar, rather than the assessment of communicative competence in authentic assessment classroom context. For instance, Adeoye et al. (2025) focused on adaptive assessment for future

learning, while Alqahtani (2021) discussed testing washback and its influence on teaching and learning. Although these studies have expanded understanding of understanding of teachers' assessment literacy, they provide limited insight into how pre-service teachers translate assessment knowledge into the evaluation of pragmatic competence during actual teaching practice. Similarly, research on pragmatic competence has predominantly examined instructional approaches or learners' pragmatic development (Azizi & Namaziandost, 2023; Fathi et al., 2025), whereas studies explicitly addressing teachers' capability to assess pragmatic competence remain scarce. Even, Choraih et al. (2016) who emphasized the importance of pragmatic competence in EFL curricula, did not investigate how teachers operationalize pragmatic competence within classroom evaluation. Furthermore, many previous investigations have relied on either quantitative surveys or qualitative case studies alone, making it difficult to capture both the level of assessment literacy and its enactment in authentic teaching contexts. Consequently, there remains limited empirical evidence explaining how pre-service EFL teachers develop and apply language assessment literacy for assessing pragmatic competence during teaching practicum, particularly in Indonesian EFL settings. Addressing this gap, the present study contributes to the literature in three important ways. First, it extends existing LAL research by examining pragmatic competence as a specific yet underexplored domain of language assessment. Second, it integrates quantitative and qualitative evidence through a exploratory sequential mixed methods design to reveal not only pre-service teachers' assessment literacy levels but also how this literacy is enacted during teaching practicum. third, the findings provide context-specific implications for strengthening practice-oriented language assessment literacy in EFL education, particularly in preparing future teachers to assess communicative language use beyond grammatical accuracy.

In response to these issues, the current study aims to investigate the language assessment literacy of pre-service EFL teachers in evaluating pragmatic ability during their teaching practicum to address the problems. The emphasis is on the application of assessment knowledge in practice, with a particular emphasis on the communicative aspects of language use that are frequently overlooked in classroom assessments. Building upon these identified gaps, this study contributes to the language assessment literacy literature by examining how pre-service EFL teachers assess pragmatic competence during teaching practicum through an explanatory sequential mixed-methods design. Unlike previous studies that primarily focused on either general assessment literacy or pragmatic instruction, this study integrates teachers' assessment knowledge, classroom assessment practices, and contextual practicum experiences to provide a more comprehensive understanding of language assessment literacy in authentic EFL settings. The findings are expected to enrich theoretical discussions on domain-specific language assessment literacy while informing teacher education programs seeking to strengthen communicative language assessment practices. Based on these assumptions, the research queries for this investigation have been set forth as follows:

1. What is the level of pre-service EFL teachers' language assessment literacy in assessing pragmatic competence?
2. How do pre-service EFL teachers assess learners' pragmatic competence during their teaching practicum?

2. Method

2.1. Research design

This study adopted a mixed-methods research design of an explanatory sequential type to examine the language assessment literacy (LAL) in pragmatic competence of pre-service EFL teachers during their teaching practicum. In an explanatory sequential design, quantitative data are collected and analyzed first followed by qualitative data, which are used to explain or elaborate the initial results (Creswell, 2014). This methodology was deemed suitable as the study sought to evaluate participants' assessment literacy and investigate its implementation in classroom practice, as well as its development through practicum experiences.

The study was executed in two successive periods. The preliminary quantitative phase aimed to create a comprehensive profile of participants' LAL concerning pragmatic capability. This phase offered a systematic summary of their knowledge, skills, and self-identified obstacles in assessment. This phase played a crucial role in mapping out their knowledge base, skills and perceived personal challenges regarding assessment. These findings proved useful for a general understanding, but did not provide sufficient insight into how assessment decisions were made in real classrooms. It is for this reason that a qualitative phase was conducted following the quantitative to explore participant assessment practices and experiences in more detail. On the other hand, the qualitative phase explored how pre-service teachers conceptualized pragmatic competence, how they operationalized these understandings into assessment tasks and criteria, and how their practices changed over time. They also specifically focused on how contextual factors, like classroom context and mentoring support, shaped their decision-making as educators. This phase offered a deeper understanding of the challenges associated with assessing pragmatics competence and strategies for doing so, drawing on participants' accounts and evidence from their classrooms.

2.2. Participants and context

The participants of this research were comprised of 68 pre-service EFL teachers who are in their sixth semester at English education department in Universitas Persatuan Guru Republik Indonesia Semarang. These students were selected because they were undertaking their teaching practicum (school-based teaching internship) at junior and senior high schools, which provided authentic opportunities for them to design, implement, and reflect on language assessment practices, including the assessment of pragmatic competence. Participants were selected through purposive sampling based on criteria relevant to the objectives of this research, i.e., (1) enrolled in teacher education program; (2) completed several courses related to language teaching and assessment methodology; and (3) currently or was undergoing a teaching practicum at schools when data collection was conducted. This context provided an opportunity for participants to perform authentic classroom assessment items that took place under mentor teachers' and university supervisors' supervision, a salient stage in the maturation of professional competencies including language assessment literacy.

2.3. Research instruments

The study utilized three principal tools to investigate various facets of pre-service teachers' language assessment literacy (LAL) on the evaluation of pragmatic competence. The utilization

of multiple instruments led to the comprehensive examination of knowledge, competence, and the actual assessment practices during the practicum context.

The first instrument was a LAL questionnaire that measured the participants' LAL level and specificity related to pragmatic competence. This questionnaire was developed based on the multidimensional LAL framework proposed by Fulcher (2012) and contextualized to the assessment of pragmatic competence. As described in the theoretical framework, the questionnaire operationalized LAL into six dimensions representing cognitive, procedural, and contextual components of assessment literacy. A total of 34 Likert scale items were grouped into six dimensions: knowledge about pragmatic competence, knowledge about assessment principles, ability to design pragmatic assessment tasks, rubric development skills, confidence in assessing pragmatic competence and perceived challenges. Items were scored on 5-point scales from strongly disagree through to strongly agree. Content validity was ensured through expert review by two professors working in areas of language assessment and pragmatics who read the instrument and provided feedback for clarity and alignment with the construct leading to a polished version of the items. Furthermore, item-total correlation analysis showed that all items positively correlated with the dimension they were intended to measure ($r > 0.30$), suggesting that the items validly represent those dimensions. Cronbach's alpha was used to establish the reliability of the questionnaire and it produced an overall coefficient of $\alpha = 0.91$, showing a high level of internal consistency (Quaigrain & Arhin, 2017). The range of reliability coefficients across the different dimensions in this study was 0.78-0.88, and supports a reliable instrument across each dimension.

The second instrument comprised classroom assessment documentation, which conveyed indicators of participants' real assessment practices during their teaching practicum. Using purposeful sampling, participants submitted a subset of documents they created and implemented in their classrooms including lesson plans, assessment tasks and scoring rubrics. With these materials, pragmatic competence was examined in the context of assessment design such as task types, features and assessment judgement criteria. The development of this instrument was validated by the expert to ensure that the analysis criteria included are representative of pragmatic competence assessment and within context. To provide measures of reliability, inter-rater agreement was determined by having a second coder who analyzed a sample of documents alongside the first and both coders assessed their agreement with one another producing strong consistency in coding (Cohen's Kappa result=0.84).

The third instrument was semi-structured interviews with a purposive subset of participants. The interviews were aimed to elicit participants' understanding on assessing pragmatic competence, their challenges in doing so and their assessment practices including how these changed during the course of the practicum. To ensure that all relevant topics were systematically covered, an interview guide was developed that still permitted probing and elaboration. To establish reliability and validity, the data were audio recorded, transcribed verbatim, and subjected to member checking; participants had an opportunity to review their responses for accuracy. These processes sought to check that data was true to the participants' perspectives and experiences.

2.4. Data collection procedures

The collection of data was carried out in multiple stages, to cover both the quantitative and qualitative components of the study, as systematically and comprehensively as possible. It starts with LAL questionnaire distribution to all participant through online (using google form). The online distribution was intended to accommodate the effective data gathering while the participants were actively conducting teaching practicum. Through this stage, the initial participants' level of LAL in relation to pragmatic competence was presented.

After the questionnaire phase, participants submitted chosen forms of practicum-directed documents such as lesson plans, assessment tasks and scoring rubrics. These materials were gathered to investigate the nature of assessment design and implementation used in real classroom settings. This stage was scheduled during a time when participants were already engaged in teaching their materials to assist in ensuring that the materials would reflect current and authentic practices instead of retrospective accounts. The collected documents served as an important source of evidence for understanding the extent to which participants' reported knowledge was reflected in their assessment practices.

Following the quantitative phase, 14 participants were purposively selected for the semi-structured interviews based on their questionnaire results and willingness to participate in the second phase of the study. The selection followed a maximum variation sampling strategy to capture diverse levels of language assessment literacy (LAL), consisting of four participants with relatively high LAL scores, six with moderate scores, and four with relatively low scores. This sampling strategy enabled the qualitative phase to explain the quantitative findings by exploring how participants with different levels of LAL understood and implemented pragmatic competence assessment during their teaching practicum. Interviews were conducted either face-to-face or via online platforms, depending on participants' availability and logistical circumstances. With participants' consent, each interview was audio-recorded for transcription purpose and content analysis.

Finally, ethical considerations were upheld throughout the data collection process. Prior to participation, all participants were informed about the study purpose and gave their consent. All data obtained from participants was completely anonymous to ensure confidentiality; all data were then combined and reported on a group basis, as participation in the study was voluntary and they could withdraw at any time without fear of penalty. This process was designed to ensure the validity of the data and protect participants' rights

2.5. Data analysis

Data analysis was conducted in accordance with the explanatory sequential mixed-methods design, whereby quantitative and qualitative data were analyzed in two related phases. The quantitative analysis was conducted first to create an overall picture of pre-service EFL teachers' language assessment literacy (LAL), and the qualitative analysis followed to give greater details on how such LAL could be enacted and developed in practice. Quantitative data from the LAL questionnaire were analyzed using descriptive statistics and SPSS 23 support. Means, standard deviations (SD), and percentages were calculated for each item and for every dimension of the instrument. Using these measures, the overall LAL level, as well as

performance in specific domains (knowledge of pragmatic competence; knowledge of assessment principles; task design; rubric development; confidence and perceived challenges) was calculated. For ease of interpretation, mean scores were classified as low (1.00–2.33), moderate (2.34–3.67) and high (3.68–5.00). Furthermore, internal consistency reliability was investigated by Cronbach's alpha coefficient to ensure the stability of the tool. The quantitative data analysis results addressed the first research question through the provision of general descriptive statistics to showcase participant level and distribution of LAL regarding assessment of pragmatic competence.

Qualitative data obtained from classroom assessment documents and semi-structured interviews were analyzed using reflexive thematic analysis following the six-phase framework proposed by Braun & Clarke (2006, 2013). This approach was selected because it provides a systematic yet flexible method for identifying, analyzing, and interpreting patterns of meaning across qualitative datasets. The analysis began with data familiarization, during which the interview recordings were transcribed verbatim and repeatedly read alongside the classroom assessment documents to develop an in-depth understanding of the data. The second phase involved generating initial codes, whereby meaningful segments related to participants' assessment knowledge, assessment practices, decision-making processes, and experiences in assessing pragmatic competence were systematically coded. In the third phase, the initial codes were examined and organized into potential themes by identifying conceptual relationships and recurring patterns across participants. These preliminary themes were subsequently reviewed and refined to ensure that they accurately represented the coded extracts as well as the entire dataset. The fifth phase involved defining and naming the themes, during which each theme was clearly described and interpreted in relation to the research questions and the language assessment literacy framework. Finally, the themes were synthesized and reported using representative excerpts from interviews and supporting evidence from classroom assessment documents to explain and elaborate the quantitative findings generated in the first phase of the study.

Following the thematic analysis, the qualitative themes were integrated with the quantitative findings during the interpretation stage using a connecting and explanatory strategy characteristic of an explanatory sequential mixed-methods design. The qualitative results were specifically used to explain the quantitative patterns observed in each dimension of Language Assessment Literacy, thereby providing a more comprehensive understanding of pre-service EFL teachers' assessment practices.

3. Findings and Discussion

3.1. Level of pre-service EFL teachers' language assessment literacy in assessing pragmatic competence

According to the quantitative results, it shows that pre-service EFL teachers below moderate levels of language assessment literacy (LAL) in assessing pragmatic competence across all measured dimensions. Although participants show a reasonable degree of knowledge about key concepts in relation to pragmatics or assessment principles, this is not straightforwardly reflected in strong pragmatic competence. Mean scores were particularly low in areas of task design, rubric development, and confidence suggesting that participants struggle to translate

their knowledge into practice or have difficulty with classroom assessment. Simultaneously, the generally high values of perceived challenges also point to the fact that pre-service EFL teachers are mindful of the richness involved in testing pragmatic competence. All these results suggest a significant disparity between conceptual understanding and practical application in particular during the teaching practicum.

Table 1. Result of descriptive statistics for pragmatic competence knowledge

Item	Statement	Mean	SD	Level
1	Understanding pragmatic competence	3.18	0.52	Moderate
2	Pragmatics and context use	3.21	0.49	Moderate
3	Knowledge of speech acts	3.09	0.51	Moderate
4	Politeness strategies	3.05	0.47	Moderate
5	Sociocultural norms	3.14	0.50	Moderate
6	Pragmatics in communicative competence	3.07	0.46	Moderate
Overall		3.12	0.48	Moderate

As depicted in Table 1, all of the items within this dimension were moderately rated with mean scores ranging from 3.05 to 3.21. Pre-service EFL teachers comprehend pragmatics in context at the highest mean ($M=3.21$), indicating a general awareness of the role situational model has on language use. Likewise, average scores on items concerning speech acts, politeness strategies or discourse and sociocultural norms imply that pre-service EFL teachers have a degree of knowledge regarding its main elements of pragmatic competence. Yet the lack of high-level responses and a fairly narrow range of scores suggests that this knowledge remains broad-stroke at best. Put differently, pre-service EFL teachers seem to recognize the terms but may not yet have a profound or functional grasp of their role in actual interpersonal communication contexts. This shortcoming hinders their capacity to measure pragmatic performance in classroom contexts.

Table 2. Result of descriptive statistics for assessment principles

Item	Statement (Shortened)	Mean	SD	Level
7	Validity	3.11	0.48	Moderate
8	Reliability	3.02	0.52	Moderate
9	Fairness	3.08	0.49	Moderate
10	Formative vs summative	3.01	0.53	Moderate
11	Alignment with objectives	3.06	0.50	Moderate
12	Appropriate methods for pragmatics	3.04	0.47	Moderate
Overall		3.05	0.50	Moderate

Similar to previous result, Table 2 mentions that pre-service EFL teachers have a fair understanding of fundamental assessment principles, with mean scores ranging from 3.01 to 3.11. Theoretically, the idea of conceptual validity, reliability and fairness are well-known as illustrated by moderately consistent scores across items. But the results also indicate that this knowledge is only abstract. In particular, although pre-service EFL teachers are knowledgeable about aligning assessment to learning objectives, there is little evidence that they can generalize this idea across the context of assessing pragmatic competence. Finally, a moderate score for choosing suitable assessment methods for pragmatics ($M = 3.04$) suggests this uncertainty also

extends to the less straightforward domain of pragmatics as compared to more general principles. Broadly, pre-service EFL teachers appear to possess procedural knowledge, but this knowledge has not yet been fully translated into context-sensitive assessment practices.

Table 3. Result of descriptive statistics for task design

Item	Statement (Shortened)	Mean	SD	Level
13	Designing pragmatic tasks	2.85	0.56	Moderate
14	Role-play activities	2.88	0.54	Moderate
15	Dialogue-based tasks	2.79	0.58	Moderate
16	Authentic interaction tasks	2.76	0.55	Moderate
17	Speech act tasks	2.80	0.53	Moderate
18	Politeness/social tasks	2.78	0.57	Moderate
Overall		2.81	0.55	Moderate

Another dimension of LAL in pragmatic competence is the seen in Table 3. It shows that the average scores of task design dimension are smaller than those presented by knowledge-based dimensions ranging from 2.76 to 2.88. The pattern indicates that the pre-service EFL teachers struggle to transition from understanding some of the concepts to applying them. While they report some ability to design tasks such as role-plays and dialogue-based activities, the moderate scores indicate that these skills are not yet well developed. Specific challenges in creating context-rich assessment contexts that reflect real communication are evident from the lower scores on designing authentic interaction tasks ($M = 2.76$). This finding also suggests that pre-service EFL teachers might be using simplified or non-authentic task formats which would not necessarily measure pragmatic competence in its entirety. In any case, the results indicate a cognitive mismatch between knowing about pragmatics and designing adequate assessment tasks.

Table 4. Result of descriptive statistics for rubric development

Item	Statement (Shortened)	Mean	SD	Level
19	Clear scoring criteria	2.68	0.58	Moderate
20	Assessing appropriateness	2.65	0.56	Moderate
21	Identifying components	2.60	0.55	Moderate
22	Speech act rubric	2.61	0.59	Moderate
23	Consistent scoring	2.58	0.60	Moderate
24	Adapting rubrics	2.66	0.54	Moderate
Overall		2.63	0.57	Moderate

Point to the dimension of rubric development, it shows one of the lowest mean scores, suggesting that rubric development is a significant area of weakness. Pre-service EFL teachers appear to struggle with defining criteria and ensuring scoring consistency, which are critical in pragmatic assessment. As shown in Table 4, rubric development is markedly one of the weakest constructs among pre-service EFL teachers with mean scores ranging from 2.58 to 2.68. All have remained within the moderate range but have persistently been near their lower limit. This indicates that, from the perspective of our pre-service EFL teachers, it is difficult to outline clear and accurate criteria for assessing pragmatic performance. Specifically, the challenges in ensuring scoring consistency ($M = 2.58$) reflect a lack of clarity on having stable evaluation

standards. In the same vein, low scores for identifying key components of pragmatic competence also imply that pre-service EFL teachers are not yet aware of a clear framework to follow when it comes to assessment. These limitations are significant because sound rubric development is necessary for a quality generalizable assessment.

The fifth dimension is the confidence in assessing pragmatic competence. This dimension presents the weakest dimension (as it scored lowest with mean (2.49 to 2.58). This implies the fact that there is a general level of unpredictability as to whether pre-service EFL teachers feel they can accurately assess pragmatic competence.

Table 5. Result of descriptive statistics for confidence in assessing pragmatic competence

Item	Statement (Shortened)	Mean	SD	Level
25	Evaluating pragmatics	2.56	0.61	Moderate
26	Assessing appropriateness	2.52	0.59	Moderate
27	Giving feedback	2.58	0.60	Moderate
28	Role-play assessment	2.49	0.62	Moderate
29	Practicum assessment	2.53	0.58	Moderate
Overall		2.54	0.60	Moderate

Table 5 shows that pre-service EFL teachers felt significantly less confident in their evaluation of role-play performance ($M = 2.49$), and the appropriateness ($M = 2.52$) indicating they might be relatively less prepared to deal with performance-measure assessment tasks. Most of their uncertainty is most likely linked to little practice and the challenges we distinguished in the previous dimensions, especially that of task design and rubric development. In classroom settings, such uncertainties could result in the avoidance of pragmatic assessment or over-reliance on simpler assessment formats. Thus, confidence is one of the elements that drives the extent to which knowledge gained through assessment translates into practice.

Table 6. Result of descriptive statistics for perceived challenges in assessing pragmatic competence

Item	Statement (Shortened)	Mean	SD	Level
30	Difficulty assessing pragmatics	3.72	0.53	High
31	Harder than grammar	3.81	0.50	High
32	Lack of training	3.76	0.52	High
33	Difficulty designing tasks	3.79	0.51	High
34	Uncertainty in scoring	3.82	0.49	High
Overall		3.78	0.51	High

In contrast to other dimensions, pre-service EFL teachers strongly acknowledge the complexity of pragmatic assessment. High scores across all items indicate that challenges are consistently perceived, particularly in task design, training, and scoring reliability. Table 6 shows that perceived challenges are consistently rated at a high level, with mean scores ranging from 3.72 to 3.82. This shows that the pre-service EFL teachers acknowledge the complexity of pragmatic competence assessment. Uncertainty in scoring ($M = 3.82$) has the highest mean score, which also indicates concerns regarding subjective and consistent evaluation. High scores on the difficulty of designing tasks and low training (two independent items) indicated

that pre-service EFL teachers feel poorly prepared to address this aspect of assessment. The perception that pragmatic assessment is more difficult than grammar-based assessment further reinforces the idea that it is seen as a demanding and less familiar domain. Instead, these findings also hint at a systematic nature of challenges occurring across different assessment aspects.

3.2. Pre-service EFL teachers' assessment of learners' pragmatic competence during teaching practicum

To answer the second research question, the qualitative phase explored how pre-service EFL teachers assessed learners' pragmatic competence during their teaching practicum. Consistent with the explanatory sequential mixed-methods design, this phase was conducted after the quantitative analysis to provide a deeper understanding of the patterns identified in the questionnaire findings. Classroom assessment documents, including lesson plans, assessment tasks, and scoring rubrics, were analyzed alongside semi-structured interviews to examine not only the assessment practices implemented by the participants but also the underlying reasoning and contextual factors that shaped those practices. Through the integrated analysis of these two qualitative data sources using reflexive thematic analysis, four interpretive themes emerged. These themes explain how pre-service EFL teachers translated their language assessment literacy into classroom practice and clarify why particular quantitative patterns, such as the moderate levels of assessment task design, rubric development, and confidence, were observed. A summary of these integrated qualitative themes is presented in Table 7, followed by a detailed discussion of each theme. The themes presented in Table 7 were derived through the thematic analysis of both classroom assessment documents and semi-structured interviews. The document analysis provided evidence of participants' actual assessment practices, including task types, assessment criteria, rubric use, and scoring procedures, whereas the interview data explained the rationale behind these practices, participants' confidence, and the contextual challenges they encountered during the teaching practicum. By integrating these two qualitative data sources, Table 7 presents a comprehensive summary of pre-service EFL teachers' pragmatic assessment practices

Table 7. Summary of interpretive themes explaining pre-service EFL teachers' pragmatic competence during teaching practicum (derived from classroom assessment documents and semi-structured interviews on pragmatic assessment practices)

Interpretive Theme	Evidence from Documents	Evidence from Interviews	Explanation of Quantitative Findings
Communicative tasks were implemented without explicit pragmatic assessment intentions	Lesson plans and speaking tasks emphasized fluency and grammar with limited reference to pragmatic objectives	Participants acknowledged uncertainty about incorporating pragmatic competence into assessment tasks	Explains the moderate score for assessment task design.

Assessment prioritized accuracy over appropriateness	decisions linguistic pragmatic	Scoring rubrics emphasized grammar, pronunciation, and vocabulary rather than contextual appropriateness	Participants perceived grammar as more objective and easier to score.	Explains why pragmatic competence was indirectly assessed.
Lack of explicit rubrics resulted in subjective judgement		Most rubrics contained holistic rather than analytic criteria for pragmatics	Participants reported uncertainty when assigning scores for pragmatic performance	Explains the low rubric development score and low confidence
Practicum shaped pragmatic assessment practices	constraints	Assessment documents reflected simplified assessment due to time and curricular constraints	Participants highlighted limited mentoring, pressure, and curriculum demands.	Explains the high level of perceived challenges

3.2.1. Communicative tasks were implemented without explicit pragmatic assessment intentions

The quantitative results showed that pre-service EFL teachers only achieved a moderate level of competence for creating pragmatic assessment tasks (overall $M = 2.81$), indicating limited ability when trying to translate their conceptual knowledge about pragmatic competence into authentic classroom assessment. This pattern is elaborated in greater detail by the qualitative analysis. Despite the fact that classroom assessment documents demonstrated the use of communicative activities, such as role-plays, dialogues, and oral presentations among participants, these activities were seldom performed with explicit pragmatic assessment goals. In contrast, the focus of both lesson plans and assessment tasks was speaking performance as a form of linguistic knowledge instead of speaking as context-sensitive language use. As a result, communicative tasks provided options for language output and not necessarily tools to assess pragmatic ability.

Likewise, analysis of classroom assessment documents illustrated that pragmatic competence was rarely a stated focus of assessment. Although most lesson plans incorporated interactive speaking activities where it was possible to elicit pragmatic performance, the learning objectives and assessment instructions rarely mentioned pragmatic features such as speech acts, politeness strategies, appropriateness, or sociocultural norms. Instead of operationalizing pragmatic competence into observable guidelines for assessment, participants seemed to take the leap that communicative activities would automatically mirror students' pragmatic ability. The result indicates that communicative task formats did not guarantee explicit assessment of pragmatic competence.

Furthermore, the interview data assist in understanding why this pattern occurred. Some of the participants shared that they chose role-play activities as task types, as these tasks were practical in nature and commonly used for speaking assessment. However, they also acknowledged that their evaluation emphasized students' linguistic accuracy and fluency over

the appropriateness of language use in different communicative contexts. This was explained by Participant 4.

I used role-play in my class, but I focused more on whether students spoke fluently and correctly, not really on whether it was appropriate. (Participant 4)

As noted by Participant 4, pre-service EFL teachers recognized the role-play was communicative, but did not feel it was a way of systematically measuring pragmatic competence. Instead, pragmatic appropriateness was considered to be an implicit byproduct of judgments rather than an explicit evaluation criterion. Thus, despite the authentic nature of communication tasks students were able to make few if any judgments about what language was socially appropriate in any given context.

The findings from the document analysis and interviews account for why in the quantitative phase participants still maintained only moderate levels of task-design competence. The issue was not that the students could not conceptualize activities to practice communicative language use, but rather concern over procedural knowledge of how pragmatic competence should be linked into these activities as an explicit construct for assessment. That is, participants knew how to teach via communication but were less sure about how to assess it pragmatically. These results indicate that the gap between conceptual change and classroom assessment practice is not in deciding which instructional activities to use but rather it is in operationalizing pragmatic competence into assessment objectives and criteria based on instruction.

3.2.2. Assessment decisions were dominated by linguistic accuracy rather than pragmatic appropriateness

The qualitative results elaborate on why participants only exhibited moderate levels of language assessment literacy revealing that assessment decisions were primarily driven by linguistic correctness rather than pragmatic appropriateness. Participants were well aware what the word pragmatic competence entailed and how important it was to use language appropriately within its context, but that was seldom taken into account when they were assessing the students. In contrast, the assessment documents paid persistent attention to linguistic features (grammar, vocabulary, pronunciation, fluency) while pragmatic aspects related to speech acts, politeness and contextual appropriateness garnered far less explicit consideration. This bias implies that participants were more confident in evaluating linguistic performance than pragmatic competence, which is contingent on the context.

Classroom assessment documents provide evidence for in this interpretation Traditional assessment criteria for speaking rubrics consisted mostly of grammatical accuracy, including the usage of vocabulary, pronunciation and fluency. Descriptors that indirectly represented pragmatic competence appeared in only a few documents, and they were frequently general descriptors like "appropriate responses" or "effective communication," without specifics about the pragmatic behaviors desired from learners. Consequently, assessment practices were mostly a replication of traditional speaking assessment rather than systematized judgment on learners' pragmatic competence. Moreover, interview findings provide additional details as to why certain decisions were made regarding these assessments. As scoring criteria were already well established for grammar and pronunciation, participant 8 consistently described himself

as more objective and easier to evaluate. In contrast, they viewed pragmatic competence as abstract and hard to evaluate in a consistent manner. Similarly, Participant 11 acknowledged that although classroom communication naturally involves pragmatic competence, assessment decisions continued to prioritize linguistic accuracy because it was considered more practical within the teaching practicum.

I am not sure about the rubric assessment criteria. I think I can just score the students' speaking based on their pronunciation. (Participant 8)

Students communicate during role-play, but when I give scores, I mostly look at grammar, vocabulary, and fluency because those are the criteria that I understand best. (Participant 11)

This finding shows that participants hold not only knowledge, but a conception on how they view the feasibility of the assessment. The preferences for linguistic features represented a desire to preserve the integrity of scoring while assuring fairness rather than an intentional neglect of pragmatic competence. Participants turned to their familiarity with observable and more measurable criteria such as grammar, pronunciation, and vocabulary even in classrooms that promoted authentic communication. To add, participants acknowledged the importance of world- or context-appropriate language use from a pedagogical standpoint, but struggled to determine adequate procedural knowledge-based assessment for functions and notional/functional grammar components translating it into an explicit criterion. Thus, assessment in the classroom was notably still focused on accuracy of linguistic features that likely explains the levels of assessment literacy being moderate throughout quantitative phase. Implications of the results for addressing the disconnect between communicative language teaching and communicative language assessment: Evidence from Chinese pre-service teacher education. These findings imply that capacity building to enhance pre-service teachers' skills in operationalizing pragmatic competence into assessable markers is crucial in narrowing the gap between communicative language teaching and communicative language assessment.

3.2.3. The absence of explicit rubrics led to subjective assessment practices

Qualitative findings also confirmed that the intermediate degree of rubric development skills observed in the quantitative phase stemmed from participants' use of general speaking rubrics rather than pragmatic competence-specific assessment criteria. While nearly all participants indicated that rubrics were vital in scoring consistently, the classroom assessment materials revealed that rubrics used during micro-teaching focused on linguistic features (grammar, vocabulary, pronunciation and fluency). Explicit tokens of pragmatic performance including speech act appropriateness, politeness strategies, sensitivity to social relationships, or contextual language use were largely left out. As a result, participants frequently relied on their own intuition in the assessment of learners' pragmatic achievement.

The rubrics used for this research were represented in most cases generic speaking assessment criteria that many teacher preparation courses introduce. Though these rubrics helped scoring for accuracy and fluency, they were of little help to judge their appropriacy in different communicative contexts. In a few lesson plans, pre-service EFL teachers expected pragmatic competence implicitly in communicative tasks such as dialogues and role-plays

without observable indicators in the scoring rubrics to reflect pragmatic behaviors. Speculatively, the fact that communicative activities were agreed as important implies participants recognized their importance for assessing pragmatic competence but did not seem to have tools or strategies by which they could evaluate pragmatic competence in a systematic manner.

This further explains why participants used such broad rubrics in the interviews. We found that participant 9 and participant 12 were unsure about how to define measurable elements of pragmatic competence, and voiced concerns regarding the reliability and fairness of scoring.

I usually use the speaking rubric provided by my lecturer because I don't know how to create specific criteria for pragmatics. If I make my own rubric, I worry that the scores will become too subjective. (Participant 9)

Grammar and pronunciation are easier because there are clear indicators. For pragmatic competence, I find it difficult to determine what deserves a high or low score, so I tend to use my overall impression. (Participant 12)

These findings suggest that reliance on holistic judgement among participants was not just a review due to lack of preparation but also a reflection of their need to salve their sense of fairness and consistency in scoring. Participants chose familiar scoring criteria that they felt more comfortable defending in the practicum setting rather than intentionally working to avoid a pragmatic assessment. But this focus on general speaking rubrics inevitably constrained their capacity to accommodate contextually appropriate language use, relegating pragmatic competence as a subtler pressure in explicit assessment of spoken language.

3.2.4. Practicum context and limited assessment literacy shaped assessment decisions

According to the qualitative findings, the structure of the individual skills or knowledge that participants possessed while judging their assessment practices was not as clear cut; rather, what they felt about conducting assessments during their practicum was also informed by contextual realities. These results provide some insight into tendency to perceive challenges during the quantitative phase. While participants acknowledged the value of assessing pragmatic competence, they often highlighted tensions between theoretical understanding gained from university and implementation in practice in classrooms. Collectively, time constraints, curriculum expectations, scant mentoring on assessment and limited opportunities for immersive exposure to pragmatics in teacher preparation influenced the design and implementation of classroom assessments among participants.

Such contextual factors were evident in the classroom assessment documents. Lesson plans tended to stress performance objectives that could be achieved in the limited instructional time available, leading to assessment activities focusing on overt forms of language use. Even though there were communicative tasks (such as role-plays and dialogues) in the classroom, assessment rubrics would become very simplified so that it was easy to score. The pattern indicates that assessment decisions were influenced by participants' assessment literacy as well as the practical considerations of pacing classroom instruction during teaching practicum. Moreover, the findings from the interview corroborate how contextual factors affected the ways those participants assessed potential students. Some of the participants expressed that they had

not received sufficient amount of pre-service training directly connected with assessing pragmatic competence and thus applied the assessment methods they had already learned or experienced.

We learned about communicative language teaching, but we were not really taught how to assess students' pragmatic competence. During practicum, I mostly followed the assessment format used by my mentor teacher because it was more practical. (Participant 5)

Sometimes I wanted to pay attention to how students used language appropriately, but with limited time and many students to assess, I focused on aspects that were easier to observe and score. (Participant 10)

According to Participant 5 and Participant 10, it demonstrates that participants' assessment decisions reflected pragmatic responses to the constraints of the practicum rather than a lack of appreciation for pragmatic competence. The reliance on familiar assessment procedures, simplified rubrics, and observable linguistic criteria represented strategies for coping with competing instructional demands while maintaining manageable and consistent classroom assessment practices. The findings also highlight the important role of mentor teachers and school assessment cultures in shaping pre-service teachers' assessment behavior during practicum, suggesting that language assessment literacy develops through interactions between individual competence and contextual opportunities for practice.

In general, the integrated analysis indicates that language assessment literacy of pre-service teachers should be positioned in a situated and developmental sense rather than as an individual body of knowledge. Participants recognized the idea of pragmatic competence and communicative assessment but felt the contextual constraints of the context made this knowledge difficult to implement in teaching practices focusing on authentic classroom assessment. These results account for the significant difficulties reported in the quantitative phase and highlight the necessity for teacher education program to include more opportunities during teaching practicum whereby material is designed, executed and evaluated using a pragmatic competence assessment approach. Strengthening cooperation between universities and practicum schools could enhance the development of assessment practices that contextualized communicative language teaching to the communicative language assessment field.

3.3. Discussion

The findings of this study indicate that pre-service EFL teachers demonstrate a moderate level of language assessment literacy (LAL) in assessing pragmatic competence, with notable variation across dimensions. To some extent, then, pre-service EFL teachers exhibit a modicum of conceptual understanding with both pragmatic competence and general principles of assessment. The findings are in line with prior research indicating that pre-service teacher education programs offer a basic understanding of assessment principles (Dreer-Goethe, 2023; Zulaiha et al., 2020). On the contrasting side, though, the findings show that such knowledge does not convert into actual expertise. Specifically, lower scores in aspects like task design, rubric development and confidence reflect that pre-service EFL teachers have difficulties

translating their knowledge into classroom practices. This gap relates to what the literature has frequently referred to as knowing about assessment versus knowing how to assess. This gap echoes the well-discussed notion in the literature of the difference between 'knowing about assessment' and 'knowing how to assess' in practice (Vogt et al., 2024). Thus, even though the pre-service EFL teachers seem theoretically prepared, their assessment literacy is partially established to apply in practice. This implies that, at least for complex language constructs like pragmatics, teacher preparation may not be adequately accommodating these needs.

Furthermore, the qualitative findings offer greater insight into the specific nature in which this gap takes place (i.e., what actually happens in classrooms). While pre-service EFL teachers predominantly used communicative tasks like role-plays and dialogues, few contained a clear focus on pragmatic competence. Instead, assessment tended to focus on linguistic features such as grammar, vocabulary, and fluency. This pattern aligns with earlier research indicating that classroom assessment often prioritizes easily measurable aspects of language over more complex and context-dependent features (Babaii & Asadnia, 2019; Kremmel & Harding, 2020). In this regard, pragmatic competence remains underrepresented not because it is considered unimportant, but because it is perceived as difficult to assess (Bouknify, 2025). Consequently, teachers might revert back to implicit or intuitive judgments instead of assessing pragmatic performance systematically. Therefore, the assessment practices do not indicate the goals of communicative language teaching. This also urges the necessity of linking communicative learning goals with those expressed at assessment in class.

In addition, this study identifies the challenges in rubric development and scoring practices as prevalent issues. Pre-service EFL teachers indicated difficulty in establishing explicit thresholds for pragmatic appropriateness leading to subjective and variable scoring. This finding accords with the notion that pragmatics assessment involves judicious evaluation and is not easily generalized to static marking rubrics (Hammouri & Al-Khanji, 2023; Korkmaz & Karatepe, 2023). Teachers can only make validity and reliability assessment decisions when rubrics are well developed. In addition, without clear criteria it is difficult to give meaningful feedback to learners which in turn limits the pedagogic value of assessment. Teachers need well-articulated rubrics to guide their assessment decisions, otherwise it is difficult to prove that they are valid and reliable (Prastikawati et al., 2025). Furthermore, the absence of specific criteria precludes quality feedback for learners, thus diminishing the pedagogical value of assessment. Furthermore, its presence is compounded by the low confidence of pre-service EFL teachers who seem to know these limitations exist without having sufficient resources to effectively remedy them. Thus, rubric development then becomes a specific area where teacher education programs need focused training (Giraldo, 2021).

Moreover, the results indicate that context impacts highly on assessment practices especially in teaching practicum. Time pressures, curricular demands, and limited mentorship support were often given as rationale for using simplified assessment methods. This phenomenon is consistent with the understanding that assessment literacy is not just an individual competency, but rather a socially- and institutionally-situated practice (Fulcher, 2012; Kim et al., 2020; Levi & Inbar-Lourie, 2020). That is, classroom environments may limit teachers' ability to enact knowledge even when they have relevant knowledge. Consequently, pre-service teachers

prioritize efficiency and practicality over sophistication which are two hallmarks of how pragmatic assessment is often neglected. This implies that LAL is not only about individual competence but also building the right conditions in which assessment takes place. Therefore, these practicum settings should be constructed to afford more structured opportunities for practicing a range of assessment approaches.

4. Conclusion

To sum up, this study indicates a moderate degree of language assessment literacy (LAL) among pre-service EFL teachers regarding measuring pragmatic competence, where conceptual knowledge outweighs procedural one. Although pre-service EFL teachers seem to have background knowledge of key concepts in pragmatics and general assessment principles, it becomes challenging for them once they are expected to create tasks or rubrics independently what should be graded and how do these reflect classroom context. It also suggests that pragmatic competence is assessed more indirectly, focusing on linguistic correctness instead of contextual appropriateness. Moreover, the two other constraints such as shortage of time and overcrowded curriculum besides inadequate training during practicum restrict enactment and development of LAL. These results suggest that current teacher education programs need to place greater emphasis on practice-oriented and domain-specific assessment training, particularly in the area of pragmatics. Yet, the limitations of this study must also be mentioned. Findings are thus limited as the data were collected from only one institution and with a small sample size. In addition, the study was based on self-reported data and intentionally chosen classroom setting that may not adequately represent assessment practices. Future research could, therefore include a large and more diverse sample size, the use of classroom observations in combination with self-reported data and/or an examination of mentoring or institutional support towards assessment literacy development. These types of studies would lead to a more in-depth insight into how pragmatic assessment (especially LAL) can be designed and implemented successfully within EFL teacher education settings.

References

- Adeoye, M. A., Prastikawati, E. F., & Riwayatiningsih, R. (2025). Beyond traditional testing: Exploring the efficacy of adaptive assessments in shaping future learners. *Journal of Nonformal Education*, 11(1), 153–169. <https://doi.org/10.15294/jone.v11i1.19934>
- Afshari, M., Beh-Afarin, S. R., & Nikoopour, J. (2024). Evaluating EAP students' pragmatic competence using authentic assessment tasks. *Issues in Language Teaching*, 13(2), 101–141.
- Alqahtani, F. (2021). The Impact of Language Testing Washback in Promoting Teaching and Learning Processes: A Theoretical Review. *English Language Teaching*, 14(7), 21. <https://doi.org/10.5539/elt.v14n7p21>
- Anam, S., & Putri, N. V. W. (2021). How literate am I about assessment: Evidence from Indonesian EFL pre-service and in-service teachers. *English Review: Journal of English Education*, 9(2), 377–388. <https://doi.org/10.25134/erjee.v9i2.4374>

- Azizi, Z., & Namaziandost, E. (2023). Implementing peer-dynamic assessment to cultivate Iranian EFL learners' interlanguage pragmatic competence: A mixed-methods approach. *International Journal of Language Testing*, 13(1), 18-43.
- Babaii, E., & Asadnia, F. (2019). A long walk to language assessment literacy: EFL teachers' reflection on language assessment research and practice. *Reflective Practice*, 20(6), 745–760. <https://doi.org/10.1080/14623943.2019.1688779>
- Bouknify, M. (2025). Task-based assessment of pragmatic competence: A moroccan EFL high school perspective. *International Online Journal of Education and Teaching (IOJET)*, 12(4), 1-15.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Braun, V., & Clarke, V. (2013). *Successful Qualitative Research a practical guide for beginners*. Sage.
- Brown, G. T. L., Andersson, C., Winberg, M., Palmberg, B., & Palm, T. (2024). Swedish teacher conceptions of assessment: A focus on improving outcomes. *Education Inquiry*, 1–21. <https://doi.org/10.1080/20004508.2024.2375111>
- Choraih, M. A., Loutfi, A., & Mansoor, A. (2016). The importance of pragmatic competence in the EFL curriculum: Application and implications. *Arab World English Journal*. <https://doi.org/10.2139/ssrn.2895549>
- Creswell, J. W. (2014). Research design: Qualitative, quantitative and mixed methods approaches (4th ed.). Sage, 12(5), 40. <https://doi.org/10.5539/elt.v12n5p40>
- Dreer-Goethe, B. (2023). Well-being and mentoring in pre-service teacher education: An integrative literature review. *International Journal of Mentoring and Coaching in Education*, 12(4), 336–349. <https://doi.org/10.1108/IJMCE-09-2022-0073>
- Esterhazy, R., De Lange, T., & Damsa, C. (2023). Performing teacher feedback literacy in peer mentoring meetings. *Assessment & Evaluation in Higher Education*, 48(2), 227-240. <https://doi.org/10.1080/02602938.2021.1980768>
- Fathi, M. J., Kafipour, R., Kashefian-Naeeni, S., & Shahsavar, Z. (2025). The impact of reflective teaching on EFL learners through implicit and explicit pragmatic competence instructions. *Reflective Practice*, 26(2), 292–310. <https://doi.org/10.1080/14623943.2024.2432063>
- Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113–132. <https://doi.org/10.1080/15434303.2011.642041>
- Fulmer, G. W., Lee, I. C. H., & Tan, K. H. K. (2015). Multi-level model of contextual factors and teachers' assessment practices: An integrative review of research. *Assessment in Education: Principles, Policy & Practice*, 22(4), 475–494. <https://doi.org/10.1080/0969594X.2015.1017445>
- Giraldo, F. (2021). Language assessment literacy and teachers' professional development: A review of the literature. *Profile: Issues in Teachers' Professional Development*, 23(2), 265–279. <https://doi.org/10.15446/profile.v23n2.90533>
- Hammouri, D. M., & Al-Khanji, R. R. (2023). Pragmatic competence in learner English: The case of Jordanian EFL learners. *Theory and Practice in Language Studies*, 13(4), 1076–1086. <https://doi.org/10.17507/tpls.1304.30>

-
- Kim, H. K., Cho, H., Yun, H., & Shin, C. (2020). Revisiting teacher-based assessment to enhance EFL teachers' assessment literacy in South Korea. *The Journal of AsiaTEFL*, 17(4), 1193–1213. <https://doi.org/10.18823/asiatefl.2020.17.4.3.1193>
- Korkmaz, S., & Karatepe, Ç. (2023). Exploring the pragmatic awareness and competence of EFL instructors at tertiary level. *Anadolu Journal of Educational Sciences International*, 13(1), 34–55. <https://doi.org/10.18039/ajesi.1086084>
- Kremmel, B., & Harding, L. (2020). Towards a comprehensive, empirical model of language assessment literacy across stakeholder groups: Developing the language assessment literacy survey. *Language Assessment Quarterly*, 17(1), 100–120. <https://doi.org/10.1080/15434303.2019.1674855>
- Levi, T., & Inbar-Lourie, O. (2020). Assessment literacy or language assessment literacy: Learning from the teachers. *Language Assessment Quarterly*, 17(2), 168–182. <https://doi.org/10.1080/15434303.2019.1692347>
- Mankki, V., Koski, J., Stenberg, K., & Poikkeus, A.-M. (2025). Teaching practicum research in Finland: A scoping review. *Educational Review*, 1–20. <https://doi.org/10.1080/00131911.2025.2506794>
- Prastikawati, E. F., Mujiyanto, J., Saleh, M., & Fitriati, S. W. (2025). Factors influencing assessment decision-making of pre-service EFL teachers: Insights from the teaching practicum experience. *Theory and Practice in Language Studies*, 15(4), 1330–1339. <https://doi.org/10.17507/tpls.1504.31>
- Quaigrain, K., & Arhin, A. K. (2017). Using reliability and item analysis to evaluate a teacher-developed test in educational measurement and evaluation. *Cogent Education*, 4(1), 1301013. <https://doi.org/10.1080/2331186X.2017.1301013>
- Tóth, E., & Csapó, B. (2022). Teachers' beliefs about assessment and accountability. *Educational Assessment, Evaluation and Accountability*, 34(4), 459–481. <https://doi.org/10.1007/s11092-022-09396-w>
- Vogt, K., Böhn, H., & Tsagari, D. (2024). Language assessment literacy. *Language Teaching*, 57 (3), 325–340. <https://doi.org/10.1017/S0261444824000090>
- Zulaiha, S., Mulyono, H., & Ambarsari, L. (2020). An investigation into EFL teachers' assessment literacy: Indonesian teachers' perceptions and classroom practice. *European Journal of Contemporary Education*, 9(1) 189–201. <https://doi.org/10.13187/ejced.2020.1.189>



