

Practices, Challenges, and Needs: Language Assessment Literacy among English Teachers in Islamic Secondary Schools

 <https://doi.org/10.30598/tahurivol19issue2page67-82>

Lutfi Indrasanti^{1*}, Fitri Rahma Wita¹

¹Universitas PGRI Sumatera Barat, Jalan Gn. Pangilun, Padang 25111, Indonesia

Abstract

This study investigates English teachers' Language Assessment Literacy (LAL) within Islamic secondary schools, focusing on MTs and MA Muhammadiyah Tanjung Bonai. In these contexts, teachers play a crucial role in evaluating students' language abilities but often face limited formal training, scarce resources, and curricula dominated by written-test orientations. This research aims to describe teachers' assessment practices, identify the challenges they face, and explore their professional development needs in LAL. Employing a qualitative multiple case study design, data were collected through semi-structured interviews with 4–6 English teachers and school principals, classroom observations, and document analysis of lesson plans, test instruments, rubrics, and assessment policies. Thematic analysis with source triangulation ensured validity. Findings reveal that teachers rely predominantly on traditional written assessments focusing on grammar and vocabulary, with limited understanding of communicative, valid, and reliable assessment principles. Key challenges include the absence of LAL-focused training, insufficient time for formative assessment, resource constraints, and tensions between national exam demands and communicative pedagogy. Despite this, teachers demonstrate growing awareness of authentic, performance-based assessment and feedback practices. The study highlights the unique Islamic school context, contributing novel insights into LAL in rural Indonesia and recommending context-sensitive professional development to enhance assessment competence and equity.

Article Info:

Keywords: Assessment Practices, Communicative Pedagogy, Islamic Schools, Language Assessment Literacy, Professional Development

***Correspondence E-Mail:**
lutfiindrsnty@gmail.com

Received manuscript: 12/02/2022

Final revision: 28/04/2022

Approved: 15/05/2022

Online Access: 20/07/2022

Published: 25/08/2022

Copyright © by the Authors

RESEARCH ARTICLE OPEN ACCESS

Publisher: Jurusan Bahasa dan Seni FKIP Universitas Pattimura, Jl. Ir. M. Putuhena, Kampus Universitas Pattimura, Poka, Ambon 97233

E-mail:
tahuri.journal@mail.unpatti.ac.id



This work is licensed under Creative Commons Attribution License 4.0 CC-BY International license

How to cite: Indrasanti, L., & Wita, F. R. (2022). Practices, Challenges, and Needs: Language Assessment Literacy among English Teachers in Islamic Secondary Schools. *Jurnal Tahuri*, 19(2), 67-82. <https://doi.org/10.30598/tahurivol22issue2page67-82>

INTRODUCTION

English language teachers in madrasah (Islamic schools) occupy a strategic position in determining the quality of language learning within Islamic educational institutions. Within Indonesia's educational system, madrasah serve a dual function: they carry the mission of both general and religious education. This duality makes language teaching and assessment practices in madrasah a compelling area of inquiry (Gudeta, 2022; Tuan, 2021). Behind the ideals of communicative pedagogy, which emphasize holistic language competence, many madrasah teachers continue to grapple with everyday realities, limited formal training in language assessment, insufficient resources, and pressure to align with the national curriculum that prioritizes written examinations. These conditions not only shape how teachers teach but also influence how they assess students' abilities, interpret learning outcomes, and design instructional follow-ups.

Language Assessment Literacy (LAL), teachers' ability to understand the principles, purposes, and practices of language assessment appropriately, is a critical determinant of instructional quality. Without adequate assessment literacy, teachers often fall into administrative routines that emphasize final scores rather than the learning process. As noted by Nugroho and Mutiaraningrum (2020) and Wijaya and Ong (2018), teachers with low levels of LAL tend to assess merely to fulfill institutional requirements rather than to support student learning. Within madrasah, this situation becomes even more complex as teachers operate within unique social, cultural, and religious contexts. While Islamic values stress fairness and honesty in assessment, the pressures of national examinations make it difficult for teachers to adopt holistic and formative approaches.

A range of studies has identified assessment literacy as a persistent challenge across educational settings. Tanjung et al. (2021) emphasize that many teachers lack formal training in designing and interpreting assessments. Chien et al. (2020) and Rasman (2018) found that English as a Foreign Language (EFL) teachers in Europe possess limited understanding of validity, reliability, and formative assessment principles. Similarly, Emilia and Hamied (2022) and Putrawan (2022) observed that teachers often rely on written tests because they are considered more practical and objective, despite their inadequacy in reflecting students' communicative competence. Research by Durriyah and Zuhdi (2018) and Sadoughi and Hejazi (2021) further revealed that language assessment remains dominated by multiple-choice and short essay tests, emphasizing grammar and vocabulary over communicative skills such as speaking and listening.

Furthermore, studies by Hussain et al. (2020) and Sasaki (2021) highlight that LAL encompasses not only theoretical knowledge of assessment but also teachers' beliefs and everyday practices. In religiously oriented schools, ethical and moral dimensions shape teachers' perceptions of fairness in assessment. Liu et al. (2021) found that teachers in culturally embedded educational contexts often face dilemmas between maintaining objective standards and adhering to social norms that emphasize harmony and respect toward students. In madrasah, such dilemmas frequently arise when teachers must balance the demands of the national curriculum with the humanistic-religious ethos that defines their institutional identity.

Several studies in Indonesia have begun to examine assessment issues within madrasah contexts. Hussain et al. (2020) and Sasaki (2021) discovered that madrasah teachers frequently modify exam formats to align with students' characteristics and Islamic values. However, such modifications are not always grounded in theoretical understanding of validity or reliability. Putra and Musigrungsi (2022) reported that most madrasah teachers have not received specialized training in language assessment, particularly in performance-based approaches that are more effective for evaluating communicative competence in real-life contexts. In contrast, Fajrina et al. (2022) showed that teachers in rural madrasah tend to rely on traditional, paper-based instruments due to limited facilities and access to digital resources.

Unlike their urban counterparts, rural madrasah teachers, such as those in Tanjung Bonai, face both structural and cultural challenges. Structurally, the lack of facilities and training opportunities constrains teachers' ability to develop LAL. Culturally, social and religious norms emphasizing modesty, obedience, and respect for teachers often make assessment practices more normative than reflective. As noted by Mubarok and Budiono (2022) and Septiana (2020), teachers' assessment practices are deeply influenced by their pedagogical beliefs and the socio-cultural contexts in which they work. Therefore, understanding assessment practices in madrasah requires close attention to local contexts, including religious values, institutional structures, and distinctive teacher–student interactions.

Although research on LAL has expanded significantly over the past decade, a gap remains in studies focusing on Islamic educational contexts, particularly at the madrasah level in Indonesia. Most existing LAL research has examined teachers in public schools, private institutions, or higher education. Studies exploring the practices, challenges, and needs of English teachers in madrasah, especially in rural areas, are still limited. Yet, these contexts embody a unique intersection of national curriculum demands, religious values, and local socio-economic realities. Kong-In and Damnet (2018) and Usman et al. (2022) argue that LAL must be understood as a contextual, not universal, phenomenon. In other words, teachers' understanding and enactment of assessment are shaped by their working environments, including cultural, policy, and experiential factors.

Against this backdrop, the present study seeks to fill a gap in existing scholarship by examining the practices, challenges, and professional needs of English teachers at MTs and MA Muhammadiyah Tanjung Bonai. By employing a qualitative multiple case study design, this research aims to capture the lived experiences of teachers rather than merely measure their knowledge levels. Through classroom observations, assessment document analyses, and in-depth interviews, the study explores assessment as both a social and pedagogical practice embedded in the everyday realities of madrasah.

The novelty of this study lies in its framing of LAL not merely as a technical competence but as a reflective practice grounded in the professional and ethical values of teachers in religious settings. By focusing on madrasah in rural contexts, this research broadens the scope of LAL scholarship, which has predominantly concentrated on urban or secular institutions. This approach enriches understanding of how teachers negotiate modern assessment demands with Islamic educational values that emphasize justice, honesty, and respect for the learning process.

Practically, this study aims to describe English language assessment practices in madrasah, identify the challenges teachers face, and reveal their perceived professional development needs in relation to LAL. Theoretically, it contributes to extending the conceptual boundaries of LAL to Islamic and rural educational contexts, areas that remain underexplored, while providing a foundation for developing more context-sensitive and equitable teacher training policies. Consequently, this research not only offers empirical insights but also opens space for reflection on the theoretical and practical development of

language assessment within madrasah education, which continues to adapt to global demands without losing its foundational values.

RESEARCH METHOD

This study employed a qualitative approach using a multiple case study design, chosen to explore in depth the experiences, perceptions, and practices of English language teachers within their everyday madrasah settings rather than to quantify variables. As Vindrola-Padros and Johnson (2020) assert, qualitative inquiry enables researchers to capture meanings constructed by individuals within specific social contexts. In this case, language assessment practices in madrasah are inseparable from the values, beliefs, and distinctive structures of Islamic education. Thus, the qualitative approach is deemed most suitable for understanding how teachers interpret and enact principles of Language Assessment Literacy (LAL) amid everyday constraints and value-laden considerations.

The research sites, MTs and MA Muhammadiyah Tanjung Bonai, located in rural Sijunjung Regency, West Sumatra, were selected purposively. The choice was based on three considerations. First, these madrasah represent typical Islamic educational institutions operating under resource constraints, both in facilities and in professional training access. Second, both institutions implement dual curricula, those of the Ministry of Religious Affairs and the national curriculum, creating particular complexities in language assessment. Third, this site has rarely been explored, offering potential for new perspectives on assessment practices in rural madrasah contexts. Purposive selection aligns with Motulsky (2021) and Ritter (2022), who contend that in qualitative research, sites and participants are chosen for their potential to yield rich, meaningful insights rather than for statistical representativeness.

Participants consisted of four English teachers (two from MTs and two from MA) and two madrasah principals involved in school-level assessment policy decisions. The teachers were selected because they actively teach English and are directly engaged in student evaluation, both in formal examinations and in-class formative assessments. The principals were included to provide institutional and policy perspectives on assessment and teacher training. Thus, the six participants offer complementary viewpoints: teachers as practitioners and principals as decision-makers. Purposive sampling was used, considering participants' direct involvement in assessment processes and willingness to engage actively in the study.

Data were collected through three primary techniques: semi-structured in-depth interviews, classroom observations, and document analysis. Interviews explored teachers' personal experiences and conceptual understandings of assessment. The semi-structured format provided flexibility, allowing researchers to adjust questions according to participants' responses (Levitt et al., 2021; Mohajan, 2018). Each interview lasted 45–60 minutes and was audio-recorded with participants' consent. Classroom observations were conducted to capture naturally occurring assessment practices, including feedback delivery, use of assessment instruments, and direct student performance evaluation. Observations were carried out multiple times at each school, supported by detailed field notes to complement interview data.

Document analysis covered lesson plans (RPP), test papers, scoring rubrics, student report cards, and institutional policies related to assessment. The purpose was to examine the extent to which LAL principles were reflected in both written and performance-based assessment practices. The triangulation of interviews, observations, and documents was applied to enhance data validity, following Khan (2019).

Data analysis followed a thematic approach based on Cho et al. (2022), comprising six iterative stages: repeated data reading, initial coding, theme identification, theme review, theme naming, and analytical reporting. This iterative process ensured accuracy and consistency of interpretation. The analysis yielded three major themes: assessment practices, challenges, and teachers' professional development needs related to LAL.

To ensure credibility and trustworthiness, the study employed both source and methodological triangulation. Source triangulation compared data from teachers, principals, and school documents, while methodological triangulation combined interviews, observations, and document analyses. Member checking was also conducted by presenting preliminary findings to participants to confirm interpretive accuracy and mutual understanding between researcher and participants. These steps were undertaken to maintain reliability and transparency while respecting participants as co-constructors of knowledge rather than mere research subjects.

RESULTS AND DISCUSSION

Overview of Teachers' Language Assessment Practices

The language assessment practices employed by English teachers at MTs and MA Muhammadiyah Tanjung Bonai reveal a dynamic interplay between the demands of the national education system, the Islamic values embedded in the school culture, and the teachers' professional capacities, which are shaped more by teaching experience than by formal training. Based on interviews, classroom observations, and document analyses, it was evident that the evaluation process remains largely dominated by conventional written tests emphasizing grammar, vocabulary, and reading comprehension. In several cases, teachers prepared multiple-choice and short-answer questions adapted from national textbooks or standardized examination packages. This pattern reflects a tendency among teachers to associate assessment with quantifiable academic performance rather than with the development of communicative competence.

In one interview, a teacher explained that at the end of each month, she was required to prepare examination questions for all classes based on a predetermined syllabus. She expressed that she had little flexibility to experiment with other assessment formats, such as presentations or oral interviews, since "the format has already been determined by the school and the madrasah supervisor." This statement indicates that assessment is perceived primarily as an administrative obligation to fulfill accreditation and reporting requirements rather than as a pedagogical tool for providing feedback to support student learning. For instance, during classroom observations in Grade VIII at the MTs, the teacher began the session with written exercises from the English in Focus textbook, followed by collective

correction. There was no reflective discussion exploring why certain answers were incorrect or how students could improve their understanding. The evaluation process was efficient but lacked opportunities for students to engage in metacognitive reflection on language as a communicative practice.

This finding aligns with the concept of Language Assessment Literacy (LAL) as defined by Genç et al. (2020), which asserts that teachers' assessment competence extends beyond technical test design skills to include cognitive, affective, and sociocultural dimensions that shape their understanding of the purposes and functions of assessment. In the context of Tanjung Bonai, teachers' understanding of assessment remains largely cognitive, focused on producing valid, reportable scores, without critical reflection on how assessment could promote meaningful learning. This was evident from one teacher, Mrs. N, who stated that "students are easier to measure through written tests because the scores are clear and can be directly recorded in the report." Such a perspective reveals a strong orientation toward administrative practicality while overlooking assessment's potential as a vehicle for communicative development.

Further analysis of the Lesson Plan (RPP) documents indicated that assessment indicators were predominantly reproductive in nature, such as "students can write sentences in the simple present tense" or "students can select the correct verb in a sentence." Rarely were indicators found that reflected communicative or contextual interaction skills. As Baker and Riches (2018) emphasize, LAL entails a reflective dimension, teachers' awareness of the social purposes of assessment within specific cultural contexts. Teachers with high LAL should be capable of aligning their assessments with community values and student needs. In the madrasah context, where honesty, responsibility, and justice are fundamental educational principles, assessment should also serve as a medium for cultivating these values rather than merely functioning as an academic screening tool.

Nonetheless, formative assessment practices within these madrasahs remain limited. Most teachers acknowledged that they rarely conducted assessments during the learning process due to time constraints and heavy administrative workloads. Another teacher, Mr. R, admitted that although he wanted to implement oral assessments through interviews or group discussions, "time usually runs out covering the material that will be tested." Observations in MA Muhammadiyah classes reflected a similar pattern: teachers managed around forty students within a ninety-minute session, leading to a focus on summative assessments such as weekly tests and midterm examinations. This pattern suggests the persistence of a summative culture, while process-oriented formative assessment has yet to be systematically integrated into classroom practice.

Resource limitations further exacerbate the issue. Support tools such as performance rubrics, observation checklists, and oral assessment instruments were nearly absent. Some teachers reported that they had to create their own assessment formats without adequate guidance. This finding underscores the lack of institutional support for developing teachers' assessment competence. According to Deygers and Malone (2019), such conditions indicate a low level of awareness and practice in LAL, teachers may recognize the importance of

alternative assessments but lack the resources and training to implement them effectively.

Within Muhammadiyah madrasahs, institutional and religious dimensions also shape assessment orientations. Teachers often emphasize fairness and honesty as moral imperatives in evaluation. One principal explained that “assessment should also reflect students’ moral character, not just their language ability.” This perspective implies that assessment is not solely a cognitive measurement but also a means of evaluating behavior consistent with Islamic ethics. However, in practice, the integration of these values into language assessment instruments remains unsystematic. Observations of rubrics revealed that affective aspects were often measured using vague criteria, such as “good,” “fair,” or “poor”, without explicit behavioral descriptors.

Overall, these findings highlight a gap between the conceptual ideal of LAL and its practical realization within madrasah contexts. Teachers exhibit strong moral awareness regarding fairness and honesty in assessment, yet lack the pedagogical tools to translate these values into valid and reliable assessment practices. As Sultana (2019) notes, LAL in Asian contexts tends to be context-dependent, meaning that assessment practices are heavily influenced by institutional structures, school culture, and community expectations. The Muhammadiyah madrasahs in Tanjung Bonai thus exemplify the intersection of Islamic identity and national education policy, producing an assessment culture that is traditional in form yet deeply moral in spirit.

Challenges in Implementing Effective Language Assessment

The challenges faced by English teachers in implementing effective language assessment at MTs and MA Muhammadiyah Tanjung Bonai reveal a multidimensional complexity that extends beyond technical issues to encompass structural, cultural, and ethical dimensions. Findings indicate that teachers operate within multiple layers of pressure: on one hand, they must comply with national curriculum standards heavily oriented toward testing; on the other, they are expected to align with the moral and spiritual values emphasized by religious educational institutions. Within this context, assessment becomes not merely a pedagogical activity but also a site of negotiation among bureaucratic demands, resource constraints, and religious values shaping everyday educational practice.

A major obstacle identified was the lack of formal training in Language Assessment Literacy (LAL). Most teachers had learned about assessment through self-directed experiences and school-based adaptation rather than through structured professional development. One teacher, Mrs. N, stated that she had never attended a specific training on language assessment in over a decade of teaching, relying instead on textbook examples and personal experience correcting student work. “So far, we just adjust to the exam format required by the madrasah,” she explained. This absence of training has directly limited teachers’ conceptual understanding of assessment, leading them to view it as an administrative procedure rather than an integral part of pedagogy.

This structural deficiency represents a key institutional barrier undermining assessment literacy, as outlined in the Teacher Assessment Literacy in Practice (TALiP) framework proposed by Berry et al. (2019). The framework suggests that assessment literacy

is shaped not only by individual competence but also by institutional support and professional development opportunities. Without such support, teachers tend to replicate traditional assessment patterns without engaging in critical reflection on their purposes or effects. This situation is further exacerbated by the scarcity of educational resources in madrasahs, ranging from teaching materials and technology to essential assessment tools such as audio devices for listening tests or performance-based rubrics.

Observations at MA Muhammadiyah revealed that classrooms relied primarily on whiteboards and textbooks as instructional tools. There were no language laboratories, and limited internet connectivity hindered teachers' access to digital assessment materials. Consequently, teachers often opted for simple written tests, which were seen as the most feasible and efficient. One teacher, Mr. R, mentioned that he wished to conduct oral tests to measure students' speaking skills but was constrained by time and facilities: "If I do an oral test, it takes too long, and the room must be quiet. Here, classes are noisy, and the schedule is tight." This statement underscores how structural limitations directly shape assessment practices, without adequate time or infrastructure, pedagogical innovation becomes nearly impossible.

In addition to these constraints, administrative workload poses a significant burden. Teachers are required to prepare extensive documentation, including test blueprints, score recaps, and progress reports, which consumes much of their instructional preparation time. Weekly observations at MTs Muhammadiyah showed that teachers often spent their breaks completing manual grade sheets in the staff room. Such administrative demands erode the opportunity to design formative or reflective assessments. As Kremmel and Harding (2020) and Levi and Inbar-Lourie (2020) note, these contextual constraints limit the enactment of teachers' assessment literacy in practice.

Another systemic pressure arises from national education policies that emphasize written test results as the main indicator of learning achievement. The English curriculum remains cognitively oriented, focusing on measurable outcomes rather than communicative competence. Consequently, teachers feel compelled to align their instruction and assessment with the national examination format. A senior teacher, Mrs. L, stated that "our ultimate goal is still the national exam, so we have to design our tests accordingly." This tension illustrates the ongoing conflict between pedagogical ideals and systemic imperatives.

Classroom observations in Grade XI at MA Muhammadiyah further confirmed this pattern: for two consecutive weeks before the semester examination, all exercises consisted of multiple-choice items and cloze tests. There were no project-based, conversational, or performance-oriented activities. Teachers were aware of these limitations but felt constrained by external demands. This aligns with Coombe et al. (2020) and Guzmán et al. (2021), who argue that in many Asian contexts, exam-oriented national policies restrict the space for formative and reflective assessment. Teachers thus operate within a teaching-to-the-test paradigm, which undermines assessment's role as a tool for learning.

A further challenge lies in the double curriculum pressure, the simultaneous demands of the national and religious curricula. Teachers in madrasahs are expected not only to teach

linguistic skills but also to integrate Islamic values such as honesty, politeness, and responsibility into the learning process. As Mr. M explained, “we want students to learn the language, but also good character. Assessment must be fair and not make students feel pressured.” This statement highlights the moral dimension of assessment practices, where teachers seek to balance academic evaluation with spiritual formation. However, dilemmas arise when these orientations conflict, for example, when institutional policies mandate minimum passing grades (KKM) while some students demonstrate stronger non-academic strengths.

Field observations revealed that the atmosphere during evaluations in madrasahs was generally calm and respectful compared to public schools. Teachers frequently allowed additional time for students who had not completed their tasks or offered remedial opportunities without penalizing scores. Such approaches reflect the internalization of rahmah (compassion) and ‘adl (justice) values within assessment practices. However, from a LAL perspective, these actions lack systematic evaluation mechanisms, such as the use of qualitative rubrics or student self-reflection. Thus, while moral principles are intuitively applied, they remain insufficiently integrated into a consistent pedagogical framework.

These findings demonstrate that teachers in madrasahs occupy a complex position: they are aware of the importance of valid, communicative, and humanistic assessment but are constrained by rigid systems and limited resources. The TALiP framework helps elucidate that such conditions stem not from individual teacher deficiencies but from insufficient contextual support, including training, policy alignment, and institutional culture, that should enable meaningful assessment practices. Hence, building sustainable LAL in madrasah settings requires structural transformation encompassing teacher training, resource provision, and the harmonization of religious values with pedagogical goals.

Teachers’ Professional Development Needs in LAL

The professional development needs of English teachers in Language Assessment Literacy (LAL) at MTs and MA Muhammadiyah Tanjung Bonai emerged as one of the most salient findings of this study. Teachers demonstrated an increasing awareness of the importance of authentic, fair, and meaningful assessment, while simultaneously recognizing their limited competence in designing and implementing such practices in the classroom. Data from interviews, classroom observations, and document analyses reveal that the desire to improve assessment quality stems not from a lack of motivation, but from insufficient access to contextualized and continuous professional learning opportunities. Teachers expressed a genuine eagerness to learn but lacked the institutional structures and professional learning models that could support a transformation of their assessment practices.

In interviews, several teachers voiced a strong desire to deepen their understanding of rubric design and performance-based assessment. Mrs. L, an English teacher at MA Muhammadiyah, explained her wish to learn how to assess students’ speaking and writing abilities without relying solely on numerical scores. She acknowledged that language ability is complex and cannot be measured merely through multiple-choice tests, yet she was uncertain where to begin in developing appropriate assessment instruments. “We know

assessment is important for learning,” she stated, “but we don’t have clear guidance on how to design communicative and fair assessments.” Such reflections reveal that teachers possess a growing awareness of the pedagogical meaning of assessment but require both conceptual and practical support to translate that awareness into classroom practice.

This awareness aligns with Qotboddin et al.’s (2020) Transformative Professional Learning theory, which asserts that sustainable pedagogical change can only occur when teachers are positioned as reflective agents rather than passive recipients of technical knowledge. According to this framework, professional learning must integrate the social, cultural, and moral contexts in which teachers work. In other words, professional development should not merely provide theoretical models or new assessment formats, but should help teachers understand how those models can be realistically implemented in religiously grounded schools with limited resources. In this sense, the transformation of LAL involves not only technical skill enhancement but also a shift in teachers’ epistemological and moral understanding of educational purpose.

Field observations revealed that teachers have begun to demonstrate small-scale innovations in assessment, even without formal support. In one MTs Muhammadiyah class, a teacher conducted a role-play activity to evaluate students’ speaking skills. Although the activity lacked a structured rubric, the teacher assessed students based on general observations of fluency and pronunciation. After the activity, she provided oral feedback to the entire class, highlighting common linguistic errors. This modest practice illustrates the potential for reflective, practice-based LAL development, provided that teachers receive relevant training and adequate institutional support.

Nevertheless, the study found a significant gap between teachers’ awareness and their actual capacity. Available training programs at the district and provincial levels remain heavily focused on curriculum administration or lesson plan design rather than language assessment. Mrs. N noted that most professional workshops she had attended concentrated on lesson planning (Rencana Pelaksanaan Pembelajaran/RPP) and compliance with national standards, with little to no emphasis on language assessment. Even when assessment topics were introduced, the focus remained on cognitive testing aligned with the national examination format rather than formative or communicative assessment. This indicates that teacher professional development programs have yet to respond to the contextual needs of English teachers in Islamic schools, who must navigate the dual demands of general curriculum requirements and institutional religious values.

This finding echoes previous studies suggesting that top-down, administrative training often fails to produce meaningful pedagogical change (Koh et al., 2017; Lam, 2019). Without reflective spaces and mentoring, teachers tend to perceive professional development as a bureaucratic obligation rather than a genuine learning opportunity. As one teacher, Mr. R, remarked, after attending several national workshops, he did not feel more prepared to conduct language assessments. “Most training only explains general theory,” he said, “but never teaches us how to design rubrics or provide effective feedback.” This reinforces the argument that LAL-focused professional learning must be practice-oriented and contextually

grounded to meet teachers' actual classroom challenges.

Structurally, the madrasah context also limits sustained professional development. There are no regular forums at the school level dedicated to discussing assessment practices. Conversations about assessment usually occur only before semester examinations or during grading meetings. As a result, assessment remains an individual practice rather than a shared pedagogical concern. In teacher rooms, observations revealed that many teachers manually scored student work without standardized rubrics, while others relied on subjective impressions. These conditions underscore the need for collective capacity building at the institutional level, not merely individual training.

From the perspective of Transformative Professional Learning theory, effective professional development should facilitate collaborative learning, critical reflection, and contextual adaptation. Teachers should be encouraged to learn from their own practice through cycles of reflection and peer dialogue. In the madrasah context, this could take the form of assessment-based lesson study, where teachers collaboratively design, implement, and evaluate authentic assessments aligned with Islamic values and students' character development. Through such a model, LAL development becomes an integral part of teachers' professional culture rather than an externally imposed project.

Moreover, training initiatives must account for the social and material realities of rural Islamic schools. Workshops on performance-based assessment, for instance, could utilize simple simulations such as peer assessment, self-assessment, or portfolio projects that require minimal technology but still promote reflective learning. Mr. M, a teacher enthusiastic about assessment innovation, noted that the most helpful training would be "something we can use directly in the classroom, not overly theoretical." This perspective highlights the need for realistic and context-sensitive approaches, training that addresses not only the "why" but also the "how" of meaningful assessment within existing constraints.

Interpreting LAL in the Islamic and Rural Context

Interpreting Language Assessment Literacy (LAL) within Islamic schools in rural areas such as MTs and MA Muhammadiyah Tanjung Bonai reveals dimensions that extend beyond technical competence. Field findings indicate that LAL in this setting is shaped by the interplay between religious beliefs, social norms, and teachers' professional responsibilities. Consistent with the concept of Situated Assessment Literacy proposed by Rad (2019), assessment literacy cannot be universally defined but must be understood as a socially situated practice embedded within the cultural and institutional contexts of teaching. In other words, assessment practices in Islamic schools are influenced by Islamic values, religious identity, and the socio-economic realities of rural communities.

In classroom practice, assessment in madrasah settings not only measures academic achievement but also aims to cultivate students' moral and spiritual character. For instance, during English classes, teachers evaluate not only reading comprehension or grammar but also honesty, discipline, and politeness during exams. One teacher, Mrs. H, stated that good assessment is not only about high scores but also about "teaching students to be honest and value their effort." This perspective positions assessment as both an academic and moral

instrument, reflecting Islamic principles such as *adl* (justice) and *amanah* (responsibility). Within this moral framework, assessment serves as a means of ethical formation as well as academic evaluation.

Tayyebi et al. (2022) argue that LAL is inherently dynamic, negotiated between professional expectations and local value systems. In Tanjung Bonai, teachers adapt modern pedagogical approaches, such as communicative language teaching, to align with the strong religious environment. Interviews revealed that teachers often modify communicative tasks to fit cultural norms. For example, during speaking assessments, teachers avoid culturally sensitive topics, opting instead for religiously appropriate themes such as daily prayers or Islamic ethics in daily life.

This process of negotiation exemplifies a form of pedagogical hybridity, blending pedagogical modernity with Islamic tradition. Yet, tensions persist. In one MA class, even when the teacher used communicative techniques, students showed greater comfort with written question–answer formats. This reflects a cultural preference for memorization and grammatical accuracy, echoing Harding and Brunfaut’s (2020) observation that such tensions represent the contextual “situatedness” of global pedagogical reforms within local educational cultures.

The rural context further shapes teachers’ understanding and enactment of assessment. The school’s geographic isolation and limited access to teacher training or educational technology compel teachers to rely on experiential knowledge and community norms. As one teacher, Mr. M, observed, “When there’s no training, we learn on our own from books or the internet, but sometimes the connection is slow.” This infrastructural limitation fosters a unique form of local professional agency, where teachers develop assessment literacy through practice, reflection, and religiously informed values rather than formal instruction.

Moreover, assessment is often perceived as a form of collective moral responsibility. Teachers consider not only students’ individual performance but also their social conduct and cooperation. In one discussion, Mrs. R noted that when assessing group work, she valued students’ ability to collaborate and respect others. For her, assessment reflects not just cognitive performance but also social harmony, a concept deeply rooted in the Islamic value of *ukhuwah* (brotherhood).

However, this moral orientation also complicates notions of objectivity. Teachers sometimes experience ethical dilemmas between maintaining fairness and showing compassion toward students from disadvantaged backgrounds. As Mr. A shared, “Sometimes we know a student is trying hard, but their score remains low. It feels wrong to give a bad grade.” This form of compassionate assessment demonstrates that fairness in Islamic schools is relational and contextual rather than strictly numerical, aligning with the situated interpretation of LAL.

The Islamic interpretation of LAL carries significant implications for language education equity in rural areas. Recognizing assessment as a socially embedded practice implies that teacher training and policy design must avoid “one-size-fits-all” approaches. As

Harding and Brunfaut (2020) emphasize, effective LAL requires teachers to critically reflect on how values and context shape assessment decisions. In the madrasah context, this means integrating Islamic values such as *adl* (justice), *rahmah* (compassion), and *amanah* (responsibility) into professional learning and assessment design.

CONCLUSION

This study concludes that English teachers' Language Assessment Literacy (LAL) in Islamic schools, specifically at MTs and MA Muhammadiyah Tanjung Bonai, represents a form of professional knowledge that develops within distinctive social, religious, and structural contexts. While assessment practices remain dominated by conventional written tests aligned with national examination formats, teachers are increasingly aware of the importance of authentic, formative, and performance-based assessment. The challenges they face, particularly limited training, time constraints, and resource shortages, reflect not personal deficiencies but systemic conditions that treat assessment primarily as an administrative duty rather than a pedagogical tool. At the same time, within Islamic education, LAL is understood as both a moral and spiritual practice grounded in justice, honesty, and compassion. Thus, language assessment in this context serves broader social functions beyond linguistic measurement. The study highlights that effective professional development in LAL must be contextualized and transformative, accounting for local values, religious culture, and rural realities. The novelty of this research lies in its interpretation of LAL not merely as a set of technical competencies but as a socially and ethically situated practice within Islamic and rural education. This perspective opens pathways for designing teacher training and assessment policies that are more equitable, reflective, and sensitive to the cultural and moral diversity of language education in Indonesia.

ETHICAL STATEMENT AND DISCLOSURE

This study was conducted in accordance with established ethical principles, including informed consent, protection of informants' confidentiality, and respect for local cultural values. Special consideration was given to participants from vulnerable groups to ensure their safety, comfort, and equal rights to participate. No external funding was received, and the authors declare no conflict of interest. All data and information presented were collected through valid research methods and have been verified to ensure their accuracy and reliability. The use of artificial intelligence (AI) was limited to technical assistance for writing and language editing, without influencing the scientific substance of the work. The authors express their gratitude to the informants for their valuable insights, and to the anonymous reviewers for their constructive feedback on an earlier version of this manuscript. The authors take full responsibility for the content and conclusions of this article.

REFERENCES

- Baker, B. A., & Riches, C. (2018). The development of EFL examinations in Haiti: Collaboration and language assessment literacy development. *Language Testing*, 35(4), 557–581. <https://doi.org/10.1177/0265532217716732>
- Berry, V., Sheehan, S., & Munro, S. (2019). What does language assessment literacy mean to teachers? *ELT Journal*, 73(2), 113–123. <https://doi.org/10.1093/elt/ccy055>
- Chien, S.-Y., Hwang, G.-J., & Jong, M. S.-Y. (2020). Effects of peer assessment within the context of spherical video-based virtual reality on EFL students' English-Speaking performance and learning perceptions. *Computers & Education*, 146(12), 103751. <https://doi.org/10.1016/j.compedu.2019.103751>
- Cho, K., Jung, K., Lee, M., Lee, Y., Park, J., & Dreamson, N. (2022). Qualitative Approaches to Evaluating Social Media Influencers: A Case-Based Literature Review. *International Journal of Electronic Commerce Studies*, 13(2), 119–136. <https://doi.org/10.7903/ijecs.2025>
- Coombe, C., Vafadar, H., & Mohebbi, H. (2020). Language assessment literacy: what do we need to learn, unlearn, and relearn? *Language Testing in Asia*, 10(1), 3–13. <https://doi.org/10.1186/s40468-020-00101-6>
- Deygers, Bart, & Malone, Margaret E. (2019). Language assessment literacy in university admission policies, or the dialogue that isn't. *Language Testing*, 36(3), 347–368. <https://doi.org/10.1177/0265532219826390>
- Durriyah, T. L., & Zuhdi, M. (2018). Digital Literacy With EFL Student Teachers: Exploring Indonesian Student Teachers' Initial Perception About Integrating Digital Technologies Into a Teaching Unit. *International Journal of Education and Literacy Studies*, 6(3), 53–60. <https://doi.org/10.7575/aiac.ijels.v.6n.3p.53>
- Emilia, E., & Hamied, F. A. (2022). Translanguaging Practices in a Tertiary EFL Context in Indonesia. *TEFLIN Journal: A Publication on the Teaching and Learning of English*, 33(1), 47. <https://doi.org/10.15639/teflinjournal.v33i1/47-74>
- Fajrina, D., Everatt, J., & Sadeghi, A. (2022). Rhetorical pattern of the Indonesian EFL undergraduate students' writings. *Studies in English Language and Education*, 9(1), 1–13. <https://doi.org/10.24815/siele.v9i1.13640>
- Genç, E., Çalışkan, H., & Yüksel, D. (2020). Language Assessment Literacy Level of EFL Teachers: A Focus on Writing and Speaking Assessment Knowledge of the Teachers. *Sakarya University Journal of Education*, 10(2), 274–291. <https://doi.org/10.19126/suje.626156>
- Gudeta, D. (2022). Professional development through reflective practice: The case of Addis Ababa secondary school EFL in-service teachers. *Cogent Education*, 9(1), 2030076. <https://doi.org/10.1080/2331186X.2022.2030076>
- Guzmán, B. L., Kouyoumdjian, C., Medrano, J. A., & Bernal, I. (2021). Community cultural wealth and immigrant Latino parents. *Journal of Latinos and Education*, 20(1), 78–92. <https://doi.org/10.1080/15348431.2018.1541801>
- Harding, L., & Brunfaut, T. (2020). Trajectories of Language Assessment Literacy in a Teacher-Researcher Partnership: Locating Elements of Praxis Through Narrative Inquiry. In M. E. Poehner & O. Inbar-Lourie (Eds.), *Toward a Reconceptualization of Second Language Classroom Assessment* (pp. 61–81). Springer International Publishing. https://doi.org/10.1007/978-3-030-35081-9_4
- Hussain, M. S., Salam, A., & Farid, A. (2020). Students' Motivation in English Language Learning (ELL): An Exploratory Study of Motivation-al Factors for EFL and ESL Adult

- Learners. *International Journal of Applied Linguistics and English Literature*, 9(4), 15–28. <https://doi.org/10.7575/aiac.ijalel.v.9n.4p.15>
- Khan, N. I. (2019). Case Study as a Method of Qualitative Research. In *Qualitative techniques for workplace data analysis* (pp. 170–196). <https://doi.org/10.4018/978-1-5225-5366-3.ch008>
- Koh, Kim, Burke, Lydia E Carol-Ann, Luke, Allan, Gong, Wengao, & Tan, Charlene. (2017). Developing the assessment literacy of teachers in Chinese language classrooms: A focus on assessment task design. *Language Teaching Research*, 22(3), 264–288. <https://doi.org/10.1177/1362168816684366>
- Kong-In, W., & Damnet, A. (2018). The Implementation of ISSECI Model for Enhancing Thai EFL Students' Intercultural Pragmatic Competence: Politeness Strategies. *Advances in Language and Literary Studies*, 9(3), 34–52. <https://doi.org/10.7575/aiac.all.v.9n.3p.34>
- 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>
- Lam, R. (2019). Teacher assessment literacy: Surveying knowledge, conceptions and practices of classroom-based writing assessment in Hong Kong. *System*, 81(1), 78–89. <https://doi.org/10.1016/j.system.2019.01.006>
- 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>
- Levitt, H. M., Morrill, Z., Collins, K. M., & Rizo, J. L. (2021). The methodological integrity of critical qualitative research: Principles to support design and research review. *Journal of Counseling Psychology*, 68(3), 357–370. <https://doi.org/10.1037/cou0000523>
- Liu, Y., Mishan, F., & Chambers, A. (2021). Investigating EFL teachers' perceptions of task-based language teaching in higher education in China. *The Language Learning Journal*, 49(2), 131–146. <https://doi.org/10.1080/09571736.2018.1465110>
- Mohajan, H. K. (2018). Qualitative research methodology in social sciences and related subjects. *Journal of Economic Development, Environment and People*, 7(1), 23. <https://doi.org/10.26458/jedep.v7i1.571>
- Motulsky, S. L. (2021). Is member checking the gold standard of quality in qualitative research? *Qualitative Psychology*, 8(3), 389–406. <https://doi.org/10.1037/qup0000215>
- Mubarok, Y., & Budiono, T. (2022). An error analysis on EFL students' writing. *Englisia: Journal of Language, Education, and Humanities*, 9(2), 187–199. <https://doi.org/10.22373/ej.v9i2.11386>
- Nugroho, A., & Mutiaraningrum, I. (2020). EFL Teachers' Beliefs and Practices About Digital Learning of English. *EduLite: Journal of English Education, Literature and Culture*, 5(2), 304–321. <https://doi.org/10.30659/e.5.2.304-321>
- Putra, J. R., & Musigrungsi, S. (2022). Indonesian EFL Teachers' Beliefs about Incorporating Cultural Aspects in Their Multilingual and Multicultural EFL Classrooms. *3L The Southeast Asian Journal of English Language Studies*, 28(1), 75–89. <https://doi.org/10.17576/3L-2022-2801-06>
- Putrawan, G. E. (2022). Translanguaging Practices in EFL Classrooms: Evidence from Indonesia. *CaLLs (Journal of Culture, Arts, Literature, and Linguistics)*, 8(1), 69. <https://doi.org/10.30872/calls.v8i1.7973>
- Qotboddin, J. M., Hossein, K., & Khalil, M. (2020). Assessment Literacy of Iranian EFL

- Teachers : A Review of Recent Studies. *The Journal of AsiaTEFL*, 17(2), 689–698. <https://doi.org/10.18823/asiatefl.2020.17.2.27.689>
- Rad, M. R. (2019). The Impact of EFL Teachers' Assessment Literacy on Their Assessment Efficiency in Classroom. *Britain International of Linguistics Arts and Education (BioLAE) Journal*, 1(1), 9–17. <https://doi.org/10.33258/biolae.v1i1.14>
- Rasman, R. (2018). To Translanguange or Not to Translanguange? The Multilingual Practice in an Indonesian EFL Classroom. *Indonesian Journal of Applied Linguistics*, 7(3), 687–694. <https://doi.org/10.17509/ijal.v7i3.9819>
- Ritter, C. S. (2022). Rethinking digital ethnography: A qualitative approach to understanding interfaces. *Qualitative Research*, 22(6), 916–932. <https://doi.org/10.1177/14687941211000540>
- Sadoughi, M., & Hejazi, S. Y. (2021). Teacher support and academic engagement among EFL learners: The role of positive academic emotions. *Studies in Educational Evaluation*, 70(12), 101–116. <https://doi.org/10.1016/j.stueduc.2021.101060>
- Sasaki, M. (2021). Effects of Study-Abroad Experiences on EFL Writers: A Multiple-Data Analysis. *The Modern Language Journal*, 91(4), 602–620. <https://doi.org/10.1111/j.1540-4781.2007.00625.x>
- Septiana, A. R. (2020). The Indonesian grammatical interference on EFL students' writing. *Journal of Research on English and Language Learning (J-REaLL)*, 1(1), 46–53. <https://doi.org/10.33474/j-reall.v1i1.5610>
- Sultana, N. (2019). Language assessment literacy: an uncharted area for the English language teachers in Bangladesh. *Language Testing in Asia*, 9(1), 1–10. <https://doi.org/10.1186/s40468-019-0077-8>
- Tanjung, F. Z., Musthafa, B., & Wirza, Y. (2021). Voice of EFL mentor teachers: Mentorship for mutual professional development. *Studies in English Language and Education*, 8(3), 986–1005. <https://doi.org/10.24815/siele.v8i3.20401>
- Tayyebi, M., Abbasabady, M. M., & Abbassian, G.-R. (2022). Examining classroom writing assessment literacy: a focus on in-service EFL teachers in Iran. *Language Testing in Asia*, 12(1), 12–24. <https://doi.org/10.1186/s40468-022-00161-w>
- Tuan, D. M. (2021). Learner Autonomy in English Language Learning: Vietnamese EFL Learners' Perceptions and Practices. *Indonesian Journal of Applied Linguistics*, 11(2), 307–317. <https://doi.org/10.17509/ijal.v11i2.29605>
- Usman, J., Zainuddin, Z., Zulfikar, T., Lugendo, D., & Yusuf, Y. (2022). When online learning and cultural values intersect: Indonesian EFL students' voices. *Issues in Educational Research*, 32(3), 1196–1212. <https://search.informit.org/doi/10.3316/informit.655017804047081>
- Vindrola-Padros, C., & Johnson, G. A. (2020). Rapid Techniques in Qualitative Research: A Critical Review of the Literature. *Qualitative Health Research*, 30(10), 1596–1604. <https://doi.org/10.1177/1049732320921835>
- Wijaya, D., & Ong, G. (2018). Applying Cognitive Linguistics to teaching English prepositions in the EFL classroom. *Indonesian Journal of Applied Linguistics*, 8(1), 1–10. <https://doi.org/10.17509/ijal.v8i1.11456>