Evaluasi: Jurnal Manajemen Pendidikan Islam
ISSN (P): 2580-3387, ISSN (E): 2615-2886
DOI : http://doi.org/10.32478/evaluasi. v9i2.0tyg8x34
Article Type : Original Research Article

Toward a Religious-Oriented Model of Speaking Assessment: Evidence from English Teachers in Islamic Secondary Schools

Siti Nurul Hidayah, Muhammad Shofiyuddin, Uswatun Hasanah

STIT Sunan Giri Trenggalek, Indonesia

Corresponding author: sitinurulhiday@gmail.com

Submission Track:

Submission : 08-09-2025

Accept Submission : 04-12-2025

Available Online : 09-12-2025

Copyright @ 2025 Author

This work is licensed under a Creative Commons Attribution-ShareAlike 4.0

Abstract.

The domain of speaking assessment has been rigorously examined within applied linguistics. Nevertheless, many contemporary investigations predominantly emphasize elements such as delivery methodologies, linguistic standards, and the validation processes of assessment instruments, frequently overlooking the incorporation of ethical and religious considerations. This research investigates the views and practices of English teachers regarding speaking assessment in Islamic high schools in East Java, aiming to develop a framework that is more attuned to cultural needs. Using a qualitative descriptive methodology, data were collected through semi-structured interviews, classroom observations, document analysis, and stimulated recall interviews with four intentionally selected teachers. The findings reveal that educators perceive the evaluation of speaking as serving a bifunctional purpose: gauging linguistic competence (encompassing aspects such as pronunciation, fluency, and grammatical accuracy) and promoting the development of students' moral and religious principles, including attributes like politeness, respectfulness, and integrity. Prevailing methodologies predominantly emphasize performance-oriented tasks, the implementation of analytic assessment rubrics, and the provision of corrective feedback; however, they infrequently integrate ethical standards cohesively and systematically. In light of these outcomes, the study puts forward the Religious-Values Infused

Speaking Assessment (RV-ISA), which aligns CEFR descriptors with Islamic ethical principles (e.g., ṣidq, ḥilm, ḥusn al-khuluq). Conceptually, this broadens the understanding of speaking assessment to encompass value-oriented factors. This model offers teachers a clear and context-sensitive rubric, enhancing communicative competence and character education within Islamic educational settings.

***Keywords**: speaking assessment, Islamic education, teachers' voices, CEFR, religious values*

Abstrak

Bidang penilaian berbicara telah dikaji secara detail dalam linguistik terapan. Namun demikian, banyak penelitian kontemporer yang dominan menekankan elemen-elemen seperti metodologi penyampaian, standar linguistik, dan proses validasi instrumen penilaian, seringkali mengabaikan pertimbangan etika dan agama. Penelitian ini menyelidiki pandangan dan praktik guru bahasa Inggris terkait penilaian berbicara di SMA Islam di Jawa Timur, dengan tujuan mengembangkan kerangka kerja yang lebih selaras dengan kebutuhan budaya. Dengan menggunakan metodologi deskriptif kualitatif, data dikumpulkan melalui wawancara semi-terstruktur, observasi kelas, analisis dokumen, dan wawancara ingatan terstimulasi dengan empat guru yang dipilih secara sengaja. Temuan ini mengungkapkan bahwa para pendidik memandang evaluasi berbicara memiliki tujuan dwifungsi: mengukur kompetensi linguistik (meliputi aspek-aspek seperti pengucapan, kelancaran, dan akurasi tata bahasa) dan mendorong pengembangan prinsip-prinsip moral dan agama siswa, termasuk atribut-atribut seperti kesopanan, rasa hormat, dan integritas. Metodologi yang berlaku umumnya menekankan tugas-tugas yang berorientasi pada kinerja, penerapan rubrik penilaian analitik, dan pemberian umpan balik korektif; namun, metodologi tersebut jarang mengintegrasikan standar-standar etika secara kohesif dan sistematis. Berdasarkan hasil ini, penelitian ini mengusulkan Penilaian Berbicara Berbasis Nilai-Nilai Keagamaan (RV-ISA), yang menyelaraskan deskriptor CEFR dengan prinsip-prinsip etika Islam (misalnya, ṣidq, ḥilm, ḥusn al-khuluq). Secara konseptual, hal ini memperluas pemahaman penilaian berbicara hingga mencakup faktor-faktor yang berorientasi pada nilai. Model ini menawarkan rubrik yang jelas dan peka terhadap konteks bagi guru, yang meningkatkan kompetensi komunikatif dan pendidikan karakter dalam lingkungan pendidikan Islam.

***Kata kunci**: penilaian berbicara, pendidikan Islam, suara guru, CEFR, nilai-nilai agama*

A. INTRODUCTION

English speaking ability has increasingly been recognized as a key competence in global communication, playing a central role in academic, professional, and intercultural domains (Dinçer & Yeşilyurt, 2013; Rao, 2019). Speaking remains particularly challenging due to limited exposure to authentic communication, teacher-centered instruction, and examination-driven curricula (Zaim, Refnaldi, & Arsyad, 2020), especially in the Indonesian context. To ensure that learning objectives are achieved, speaking assessment is a measure of oral performance and a pedagogical tool to provide feedback and guide instructional practices (Dao & Nguyen, 2020; Duque-Aguilar, 2021).

In second language acquisition, speaking competence is one of the most important domains, as it helps learners demonstrate linguistic competence and engage in meaningful communication in various branches, such as academic, social, and professional domains (Hughes, 2003; Luoma, 2004). Speaking competence is challenging to teach and assess in the context of Indonesian English teaching. Most classes are large classes with minimal exposure to authentic communication in English. The grammar-translation tradition still dominates the class, restricting opportunities for students to develop proficiency in the language (Zaim, Refnaldi, & Arsyad, 2020; Richards, 2015). Consequently, English teachers need to design purposeful assessment practices. The evaluation process must enable educators to document communicative proficiency and furnish constructive feedback (Fulcher, 2010; Brown & Abeywickrama, 2018). Constructive feedback is crucial for educators to assess learners' linguistic capabilities, appraise their language competencies, and identify challenges that arise within the educational environment (Dao & Nguyen, 2020).

English teaching practice in Islamic High Schools in Indonesia urges an effective speaking assessment. On its unique curriculum and pedagogy, Islamic schools are integrated with strong Islamic values and characters, which opens opportunities to focus on operationalizing speaking assessment integrated with Islamic values and characters. This practice follows the Ministry of Religious Affairs, emphasizing the importance of integrating moderation values, including in English language teaching (Azra, 2018; Kementerian Agama RI, 2020). Not only as a technical act of quantifying linguistic performances, this assessment of English-speaking proficiency will also act as a pedagogical opportunity to reinforce Islamic values, including honesty, respect, humility, and responsibility (Halstead, 2004; Arthur, 2008). Integrating English education with Islamic values places unique demands on teachers and raises critical issues regarding how speaking assessment is conceptualized and implemented in this context. In addressing these challenges, extensive exposure to vocabulary plays a crucial role, as it enhances students' speaking ability, directly supports both speaking practice and assessment, and increases motivation with games and other interactive methods (Hasanah & Nisyak, 2025).

The previous studies in the last few years have focused on delivery modes, technology-mediated testing, traditional linguistic criteria, and the validation of scoring instruments rather than just the technical and linguistic dimensions of performance measurement (Guo, 2024). Brown (2006) investigates the IELTS speaking test, in which

fluency, coherence, pronunciation, and accuracy are the primary indicators for raters' evaluation of oral performance. Fulcher (2010) focuses on a comprehensive framework for reliability and fairness of language testing. There is an urgent need to explore how speaking assessment might relate to broader socio-cultural and ethical dimensions in an Islamic educational setting to integrate the religious values with learning objectives. On this matter, socio-cultural and moral considerations might shape assessment practices, with explicit measurement and concepts, especially in the Islamic school context.

The study implements a novel framework for speaking assessment in Islamic educational contexts by developing and validating the Religious-Values–Infused Speaking Assessment (RV-ISA) rubric. Unlike conventional instruments that focus exclusively on linguistic performance, the RV-ISA rubric aligns core descriptors from the CEFR—such as fluency, interaction, discourse management, and pragmatic competence—with key Islamic values. These include *ṣidq* (truthfulness in content), *ḥilm* (patience in turn-taking), *tawāḍu'* (humility in responding to arguments), *iḥsān* (carefulness and academic responsibility), and *ḥusn al-khuluq* (politeness in diction and non-verbal behavior). Integrating these dimensions reflects Islamic education's ethical and character-based foundations and resonates with broader discussions in moral education (Arthur, 2008; Halstead, 2004).

Religious principles are interwoven into measurable criteria within oral assignments, including problem-solving dialogues, simulated roles, and debates. Along with the established procedure, an expert reviews the instrument's validity. This study employs the importance of teachers' voices in co-designing and testing the rubric. The teachers' role in evaluating the usability reflects the realities of assessment practice in Islamic High Schools. It corresponds with language assessment literacy (LAL) on empowering teachers to integrate socio-cultural and ethical dimensions into assessment implementation (Inbar-Lourie, 2013; Vogt & Tsagari, 2014; Xu & Brown, 2016). Communicative competence and moral development are expected to generate positive teaching and learning in English-speaking skills.

This study addresses a critical conceptual and empirical gap in language assessment by involving religious values, which are operationally defined and empirically validated. Aligning Islamic values in the instrument through the RV-ISA rubric allows the study to expand the scope of assessment research. Socio-cultural and moral constructs should improve the theoretical discussion on the multidimensionality of speaking proficiency (Zhu, Mofreh, & Salem, 2023). Practically, the RV-ISA rubric explores the transparency and readiness of the assessment tool for teachers that correlates linguistic performance and moral conduct. This model resonates with the national purpose of education in madrasah by strengthening knowledge and character (Kemenang Agama RI, 2020).

Building upon the conceptual and practical issues outlined above, this study is guided by research questions designed to illuminate current practices and innovative possibilities in speaking assessment within Islamic high schools. The study aims to explore how English teachers perceive and carry out speaking assessments in their classrooms, and how these approaches either incorporate or ignore the incorporation

of ethical and religious principles. In addressing this issue, the study further asks how linguistic components can be systematically aligned with Islamic teaching, such as *ṣidq* (truthfulness), *ḥilm* (patience), *tawāḍu'* (humility), *iḥsān* (academic responsibility), and *ḥusn al-khuluq* (politeness), to construct a linguistically rigorous and value-oriented rubric. A central question concerns the extent to which the resulting Religious-Values–Infused Speaking Assessment (RV-ISA) rubric demonstrates validity, reliability, and usability across multiple schools. Finally, the research examines teachers' perceptions of the relevance, practicality, and pedagogical impact of employing such a rubric, thereby shedding light on its potential contribution to language assessment literacy and character-based education in the madrasah context.

B. METHODOLOGY

This research utilized a qualitative descriptive methodology (Merriam & Tisdell, 2016) to investigate educators' experiences, beliefs, and pedagogical practices within a naturalistic context. This design supports the research to capture the complexity of the teacher's conceptualizations in conducting speaking assessment and the contextual factors developing the application of the assessment (Creswell & Poth, 2018). The study documented the teacher's practices in assessment and employed the insights to co-construct and validate the RV-ISA rubric. This combination is oriented to the exploratory and developmental process of the assessment.

Purposive criterion sampling was implemented to select the participants (Patton, 2015), emphasizing high-performing English teachers in Islamic High School across East Java. The criteria of "high-performing" involve teachers who were at least one or two of the following:

1. Coaching or mentoring students for district, provincial, or national speech/debate competitions
2. Receiving pedagogical awards or official recognition
3. Demonstrating consistent student improvement in speaking scores
4. Being recommended by the school principal as exemplary assessors
5. Having at least five years of teaching experience supported by assessment records.

The participating teachers and schools were:

1. Ahmad Zubaidi Amrullah (male) is an English teacher at MAN 2 Gresik. His school has won multiple national competitions, including first place in the Institute English Competition (Institut Pendidikan Tapanuli Selatan, 2021) and awards in international English contests organized by UIN Antasari.
2. Sari Wulandari (female), English teacher at SMP Islam Baitul Izzah, Nganjuk. The school won second and favorite prizes at the SMASA Virtual Olympiad Competition (SMAN 1 Nganjuk, January 2022).
3. Sayidi Hasbullah (male), English teacher at SMP Alajidiyah Palduding, Pamekasan. His school was the first winner of the English Festival Speech Contest (University of Islam Madura, 2020).

4. Rizky Dian (female) is an English teacher at SMK Al Islah, Surabaya. Her school won first place in the Speech Olympics held by the Sidoarjo Education Office.

The rationale for choosing these schools and teachers was to capture the voices of practitioners who had demonstrated excellence in English language teaching and speaking competitions, thereby ensuring that the study documents best practices.

The data were collected to apply methodological triangulation (Denzin, 2012), thereby minimizing self-report bias and enabling the observation of conformity of the teacher's beliefs and practices. The data collection proceeded in four stages:

1. Initial interviews – Each teacher participated in a semi-structured interview to explore their teaching background, experiences in speaking assessment, and perspectives on integrating religious values into their classroom practices.
2. Classroom observations with video recording—Teachers were observed during speaking assessment sessions. Video recordings were taken to support the validity of the first interview data and provide concrete evidence of classroom practices.
3. Stimulated recall and post-observation interviews—After the observations, Teachers were shown selected excerpts from their classroom videos and asked to reflect on their assessment practices and whether they aligned with the RV-ISA rubric (Rubric of Values-integrated Speaking Assessment). This step allowed participants to examine their practices critically and provided deeper insight into how assessment and values integration were enacted.
4. Document collection – Supporting documents (lesson plans, assessment rubrics, and student performance records) were gathered to triangulate findings from interviews and observations. These documents served as additional evidence for validating the themes identified.

Firstly, semi-structured interviews collected the primary data, conducted online via Zoom or phone calls in November 2024, each lasting 45–60 minutes. The protocol consisted of seven adapted questions (Duque-Aguilar, 2019; Kusumawati & Sumardi, 2020; Köroğlu, 2021) covering topics such as the definition, importance, instruments, focus, and feedback in speaking assessment, as well as recommendations for improvement. All sessions were audio-recorded, transcribed verbatim, and member-checked with participants for accuracy.

Secondly, classroom observations were conducted four times in each participating school with structured checklists and field notes. The subsequent step involved the collection and thorough review of various documents utilized in the classroom setting, including: assessment rubrics, lesson plans, speaking test scripts, and teacher feedback samples. Stimulated recall interviews were conducted by replaying selected classroom interactions with participants, prompting them to reflect on and explain their assessment decisions. This approach allowed access to teachers' tacit reasoning and decision-making processes during assessment.

Data were analyzed using a qualitative descriptive approach (Sandelowski, 2000), emphasizing low-inference interpretations of participants' accounts. Rather than using specific coding techniques, the analysis focused on summarizing the participants' answers to each interview question in a matrix format, allowing for comparisons across cases and identifying commonalities and differences. This descriptive synthesis was further reinforced by observations and document analysis notes, allowing the researchers to present the findings in the participants' language while maintaining the richness of the context. Member checking and triangulation were employed to improve the descriptive validity of the analysis.

C. RESULT AND DISCUSSION

The results of this research elucidate a complex representation of oral assessment methodologies in Islamic secondary educational institutions within East Java. Although educators predominantly utilized traditional methodologies such as performance-oriented tasks, analytic evaluation frameworks, and assessments emphasizing linguistic precision, they also exhibited an implicit inclination towards incorporating ethical and religious principles within their evaluative practices. This dualistic approach exemplifies adherence to international tendencies in language assessment and the emergence of localized adaptations influenced by the foundational principles of Islamic pedagogy. In line with Creswell and Creswell's (2018) assertion that discussion sections should connect results to broader theoretical and practical issues, the following discussion synthesizes the seven identified themes by relating teachers' voices to established frameworks such as the CEFR (Council of Europe, 2020), research on classroom-based assessment (Black & Wiliam, 2018), and the growing literature on assessment literacy (Scarino, 2013; Xu & Brown, 2016). This approach allows us to highlight not only the alignment and divergence between local practices and international standards but also the unique contribution of this study—the proposal of a Religious-Values Infused Speaking Assessment (RV-ISA). The following is the RV-ISA rubric, which was adapted and subsequently validated by field experts. The content validation involved Dr. Siti Nurul Hidayah, M.Ed., from STIT Sunan Giri Trenggalek (language assessment expert), and Dr. Afrizal El Adzim Syah Putra, M.A., from UIN Sayyid Ali Rahmatullah Tulungagung (Islamic studies expert). This rubric was validated through expert review and implemented in the present study. The final version is presented below:

Table 1. RV-ISA Rubric for Speaking Assessment

Dimension	CEFR-Based Indicator	Islamic Value Indicator	Score 1 (Beginning)	Score 2 (Developing)	Score 3 (Proficient)	Score 4 (Advanced)
Fluency	Flow of speech, pauses	—	Speech is fragmented, with frequent	Speech flows with many pauses	Mostly smooth with minor pauses	Smooth, natural, sustained

			breakdowns.			
Pronunciation & Intelligibility	Accurate articulation, comprehensibility	—	Frequently unintelligible, heavy influence of L1	Sometimes unclear, intelligible with effort	Generally clear, occasional mispronunciation	Very clear, natural pronunciation
Interaction	Turn-taking, responsiveness	ḥilm (patience, respect)	Interrupts, ignores cues	Sometimes interrupts, limited patience	Generally polite, waits for their turn	Always patient, encourages peers
Content & Accuracy	Relevance, factual correctness	ṣidq (truthfulness)	Content irrelevant, many false statements	Some relevant points, occasional inaccuracies	Mostly relevant, accurate content	Highly relevant, fully accurate
Politeness & Attitude	Diction, manner of speaking	ḥusn al-khuluq (good conduct)	Disrespectful, impolite expressions	Sometimes polite, inconsistent	Polite and respectful overall	Exemplary politeness, respectful tone
Humility & Cooperation	Acknowledging others, teamwork	tawāḍuʿ (humility), iḥsān (excellence in doing good)	Shows arrogance, dismissive	Occasionally arrogant, limited cooperation	Cooperative, acknowledges peers	Very humble, supports and uplifts others

Having established and validated the RV-ISA rubric, the instrument was subsequently implemented in classroom observations and teacher reflections to capture how linguistic and value-based indicators were enacted in practice. The data obtained from interviews, classroom observations, stimulated recall sessions, and document analysis were analyzed thematically. This process generated seven major themes that illustrate how teachers integrated both CEFR-aligned speaking criteria and Islamic values into their assessment practices. The following section presents these themes, supported by excerpts from teacher interviews and classroom evidence.

Theme 1. Conceptualizations of Speaking Assessment

The teachers in this study demonstrated diverse conceptualizations of speaking assessment. One teacher emphasized its function as a tool to measure linguistic accuracy and performance: “Speaking assessment is to test students’ ability in pronouncing words correctly and fluently” (Teacher A). Others viewed it more broadly as part of a holistic evaluation process, not limited to linguistic forms but also interaction and character: “It is not only about grammar and fluency but also how students behave and respect each other when speaking” (Teacher B).

This variety reflects what Hughes (2003) and Luoma (2004) describe as the ongoing debate between product-oriented and process-oriented perspectives in oral assessment. While some teachers equated speaking assessment with testing

pronunciation and fluency—an orientation typical in speech contests—others were closer to communicative and socio-pragmatic approaches. Recent literature has also noted this divergence. Teachers' conceptualizations of assessment often oscillate between testing discrete linguistic skills and evaluating communicative competence in authentic contexts (Cumming, 2009; Yuzar, 2020).

The findings suggest that in the context of Islamic schools, teachers' conceptualizations already hint at integrating ethical or religious dimensions, even if implicitly. This extends beyond conventional linguistic descriptors in CEFR (Council of Europe, 2020) and aligns with calls for more contextualized frameworks of assessment that capture local educational values (Rohmah & Hanifiyah, 2023; Muttaqien, 2017).

Theme 2. Purposes and Importance of Speaking Assessment

The oral proficiency evaluation was deemed significant among all respondents, although the articulated objectives differed. One educator highlighted its significance in equipping learners for competitive events: "The assessment of speaking skills is essential, particularly given that our institution participates in numerous oratory and narrative competitions" (Teacher C). Another highlighted communicative competence as the central aim: "The purpose is to make students confident to speak English in real life, not only for exams" (Teacher D).

These responses resonate with prior studies reporting that speaking assessment is often shaped by institutional demands and cultural expectations. Dao and Nguyen (2020) showed that speaking assessment was closely tied to communicative competence within the Communicative Language Teaching (CLT) paradigm in Vietnamese EFL contexts. Similarly, in Indonesia, Zaim et al. (2017) found that teachers valued speaking assessment for academic accountability and for fostering real-world communication skills. Evaluating speaking practice within the classroom is crucial for assessing the effectiveness of instructional methodologies employed by educators and understanding their impact on learners. This importance is supported by research conducted by Hidayah (2022) and Hidayah & Hasanah (2020).

However, the dual orientation distinguishes the current context: while teachers acknowledged the global communicative purposes of speaking assessment, they implicitly associated it with moral and religious formation. This exemplifies the overarching educational objective of madrasah, wherein evaluation is anticipated to function not solely as an indicator of scholarly advancement but also as an enhancement of adab and moral character (Kementerian Agama RI, 2020). By positioning speaking assessment at the intersection of linguistic, pragmatic, and religious purposes, teachers in Islamic schools provide a nuanced rationale that has been underexplored in the literature.

Theme 3. Performances and Classroom Tasks

Educators articulated using diverse classroom assignments to evaluate students' oral communication abilities. Frequently cited endeavors encompassed role-playing exercises, narrative construction, formal debates, and oratory presentations, frequently designed to emulate competitive formats. One teacher noted: *"We usually assess students by asking them to deliver a speech or storytelling, because these are also part of English competitions"* (Teacher A). Another highlighted dialogic task is: *"Sometimes I give students role-play or debates so they can practice interaction and not just memorization"* (Teacher B).

This emphasis on performance-based tasks aligns with previous findings in EFL contexts where role-play, presentations, and debates are widely used as assessment formats (Christie & Listyani, 2018; Duque-Aguilar, 2021). Such assignments are esteemed due to their ability to replicate genuine communicative occurrences and enable learners to exhibit fluency, interaction, and discourse management. However, the findings also reveal a strong alignment with competition culture, especially in schools that actively participate in regional or national contests. While this may cultivate motivation and increase visibility, it concurrently poses the danger of constricting evaluative practices to a restricted array of high-stakes performances.

At the same time, teachers in Islamic schools incorporated religiously informed tasks in subtle ways. For instance, storytelling often involved narratives from Islamic history or moral parables, linking linguistic practice with character formation. This provides evidence that speaking assessment in this context already integrates religious content, even if the rubrics remain primarily linguistic. Incorporating content imbued with values exemplifies what Scarino (2013) articulates as the interconnection among language, culture, and identity within evaluative methodologies.

Overall, while the classroom tasks observed are consistent with established communicative assessment techniques, their adaptation to Islamic educational settings creates opportunities for the Religious-Values Infused Speaking Assessment (RV-ISA) model. Specifically, familiar tasks such as debates and storytelling can be enriched with descriptors that measure honesty (*ṣidq*), patience (*ḥilm*), and politeness (*ḥusn al-khuluq*) alongside fluency and accuracy. This both contextualizes the assessment within local educational aims and contributes a novel dimension to the broader field of speaking assessment.

Theme 4. Assessment Instruments and Approaches

Educators indicated using various methodologies and instruments to assess students' verbal proficiency. The tools most commonly cited included analytic rubrics and observation checklists; however, certain educators also depended on more subjective evaluations. One teacher stated: "I usually use a rubric to check pronunciation, fluency, and grammar, but sometimes I just give scores based on overall impression" (Teacher

C). Another teacher emphasized communicative approaches: “I try to see how students interact with each other, not only their accuracy” (Teacher D).

These methodologies correspond with international scholarship in which rubrics are predominant in assessing speaking abilities (Brookhart, 2018; Venera, 2017). People often commend analytic rubrics for their role in boosting clarity and fairness, which helps make assessment criteria understandable and uniform (Jonsson & Svingby, 2007). Nonetheless, empirical investigations warn that rubrics may evolve into formulaic constructs, emphasizing linguistic attributes while overlooking interactional or socio-pragmatic aspects (Fulcher, 2015).

The results obtained from Islamic educational institutions elucidate this dichotomy. While educators acknowledged the value of rubrics, their descriptors were largely restricted to conventional criteria such as pronunciation, grammar, vocabulary, and fluency. Socio-cultural or ethical dimensions were seldom considered. This phenomenon corresponds with what Ojochegbe (2024) elucidates as the shortcoming embedded in 'standardized' rubrics, which frequently do not encompass the full complexity of communicative competence within contextualized environments.

Importantly, teachers in this study expressed a desire for more contextually relevant instruments. One teacher noted: “We need rubrics that include attitude and politeness, because that is part of our school’s vision” (Teacher B). This observation underscores a central contribution of the present research: the possibility of designing a Religious-Values Infused Speaking Assessment (RV-ISA) instrument. By aligning CEFR descriptors (fluency, interaction, discourse management) with Islamic values such as *ṣidq* (truthfulness) and *ḥusn al-khuluq* (politeness), such an instrument would address both the fairness expected of rubrics and the moral-ethical mission of Islamic education.

Consequently, while educators currently rely on conventional rubrics and subjective assessments, the research highlights the limitations of existing methodologies and the potential for progress through instruments specifically designed for particular contexts. This corresponds with wider appeals for assessment literacy encompassing technical proficiency, ethical considerations, and cultural responsiveness (Scarino, 2013; Xu & Brown, 2016).

Theme 5. Focus of Assessment Criteria

The findings revealed that teachers primarily focused on pronunciation, fluency, and grammar when assessing students’ speaking skills. One teacher explained: “The first thing I notice is pronunciation and fluency, because if students are not fluent, the message will not be clear” (Teacher A). Another added: “Grammar is also very important, since mistakes can change meaning” (Teacher C). These emphases reflect a traditional orientation to linguistic accuracy, which continues to dominate many EFL classrooms.

Though these emphasized areas match current linguistic specifications (Luoma, 2004; Hughes, 2003), they suggest a shortcoming in recognizing speaking proficiency. Preceding evaluations have indicated that communicative proficiency comprises linguistic resources and a command of pragmatic and relational dimensions (Council of Europe, 2020). Educators may unintentionally neglect other equally critical elements by emphasizing pronunciation and grammatical accuracy, such as facilitating interactions, articulating ideas coherently, or employing language appropriately within social contexts.

Interestingly, some teachers indicated awareness of broader dimensions. One remarked: “I also pay attention to how students show respect to their friends while speaking” (Teacher B). Such attention to values suggests an emerging orientation toward integrating ethical and religious aspects alongside linguistic features. This is noteworthy given the context of Islamic schools, where education is expected to nurture communicative ability and character (Kementerian Agama RI, 2020).

The gap between traditional language standards and the unspoken acknowledgment of principles highlights the shortcomings of modern assessment techniques. Academic professionals persist in employing time-honored standards from competitions and evaluations; however, they simultaneously admit the relevance of courtesy, sincerity, and humility in judging students' performance. This dual focus points to the potential of developing a Religious-Values Infused Speaking Assessment (RV-ISA) rubric that formally incorporates ethical indicators—such as ṣidq (truthfulness) and ḥilm (patience)—alongside CEFR descriptors.

Therefore, while the teachers’ current practices mirror mainstream approaches to oral assessment, the findings highlight a limitation (linguistic reductionism) and a novel opportunity: enriching assessment criteria with culturally and religiously grounded values. This would align with international frameworks of communicative competence and advance a more holistic, contextually responsive assessment model in Islamic high schools (Scarino, 2013; Xu & Brown, 2016).

Theme 6. Feedback Practices

Teachers in this study reported providing feedback to students in multiple forms, ranging from immediate oral corrections during classroom activities to delayed written or oral comments after assessment tasks. One teacher explained: “I give direct correction when students make mistakes in pronunciation, so they can immediately know the right way” (Teacher A). Another highlighted delayed feedback: “Sometimes I give feedback after the performance is over, because I don’t want to disturb their confidence” (Teacher D).

This collection of methodologies corresponds with more comprehensive findings documented in the academic literature. Feedback is extensively acknowledged as a critical element of formative assessment, as it directs learners towards enhancing their

performance and fosters learner autonomy (Hattie & Timperley, 2007; Gamlo, 2019). Immediate feedback effectively addresses discrete linguistic errors, while delayed feedback encourages reflection and reduces anxiety (Sakale, 2019).

The findings also show that feedback was not limited to linguistic aspects. One teacher remarked: “I remind students to be polite in their language and behavior, not only correct in grammar” (Teacher B). This demonstrates how feedback in Islamic school contexts often extends to moral and behavioral guidance, consistent with the schools’ broader character education mission. Such practices align with the growing recognition that language assessment should encompass linguistic performance and socio-cultural values (Scarino, 2013).

Nevertheless, the feedback practices observed also reveal gaps and challenges. Teachers reported limited time to provide individualized feedback, and some relied heavily on corrective feedback without offering constructive strategies for improvement. Moreover, while teachers occasionally addressed the concepts of politeness and respect, these elements were not consistently incorporated into feedback frameworks. As a result, corrective feedback may lead to misunderstandings and conflict in the classroom, as it is often perceived as negative criticism regarding students' mistakes. As Hidayah (2016) suggests, positive reflective comments and encouraging feedback from teachers can motivate students to be more forthcoming about their learning challenges, facilitate the exchange of reflective ideas, and foster a more meaningful learning experience with their teachers. This highlights the necessity for structured rubrics or feedback protocols that explicitly include linguistic and value-based criteria to foster a more positive learning environment in the classroom.

The study, therefore, highlights the potential of the Religious-Values Infused Speaking Assessment (RV-ISA) model to formalize feedback practices. By including descriptors for fluency and ethical conduct, RV-ISA could guide teachers to provide feedback that is not only corrective but also developmental, reinforcing both communicative competence and moral values. In this manner, feedback emerges as a conduit for linguistic acquisition and the cultivation of Islamic values, thereby corresponding with the objectives of madrasah education (Kementerian Agama RI, 2020).

Theme 7. Teachers’ Recommendations and Aspirations

Teachers also shared several recommendations for improving speaking assessment in their contexts. Many emphasized the importance of adequate preparation and scaffolding to support students before assessment tasks. One teacher suggested: “Students should be given time to practice and prepare vocabulary before the assessment, so they don’t just memorize but understand” (Teacher C). Another highlighted the role of school-wide initiatives: “It would be good to have English Days or regular speaking activities, not only when there is a test” (Teacher D).

These recommendations echo broader calls in the literature for embedding speaking assessment within everyday classroom practices rather than treating it as isolated, high-stakes events (Davison & Leung, 2009; McNamara, 2019). Utilizing assessment strategies within regular educational frameworks allows instructors to alleviate test-related stress and support a more accurate reflection of communication proficiency. Furthermore, expanding students' vocabularies is crucial for fostering a more interactive and positive classroom environment (Hasanah & Hidayah, 2025). A robust vocabulary enhances students' comprehension and enables them to articulate their ideas clearly and communicate effectively across diverse contexts.

Aspirations for more meaningful and enjoyable assessments also emerged. Teachers desired assessment tasks that motivate students, reflect real-world communication, and incorporate cultural and religious content. As one teacher put it: “Speaking assessment should not only be about marks. It should also help students to be polite, respectful, and confident” (Teacher B). This objective aligns with scholarly literature that promotes the concept of assessment for learning instead of merely assessment of learning (Black & Wiliam, 2018).

Importantly, these recommendations underscore teachers’ awareness of the limitations of current practices—particularly the dominance of competition-style performances—and their interest in assessment models that align better with Islamic schools' character-building mission. The proposed Religious-Values Infused Speaking Assessment (RV-ISA) responds directly to these aspirations by operationalizing values such as ṣidq (truthfulness) and ḥusn al-khuluq (politeness) alongside linguistic descriptors.

The teachers’ recommendations reflect practical concerns (preparation, frequency, task variety) and aspirational goals (integration of values, meaningful learning). This dual orientation strengthens the argument that speaking assessment in Islamic schools should evolve toward a more holistic framework, simultaneously promoting communicative competence and moral character.

The findings across the seven themes elucidate a multifaceted landscape of speaking assessment methodologies within Islamic secondary educational institutions. The way educators see assessment goes beyond just measuring language capability; it also serves as a tool for instilling ethical and moral values. Although their existing methodologies resonate with contemporary global trends in performance-based assessment, rubrics, and feedback mechanisms, these approaches predominantly adhere to conventional linguistic standards. Concurrently, implicit and explicit allusions to religious principles reveal a distinctive opportunity: construct assessment frameworks that are methodologically robust and ethically sound. This research tackles a conceptual and practical void in existing literature by introducing the RV-ISA model. It broadens the focus of speaking evaluation to include fluency, accuracy, discourse management, and values fundamental to Islamic education. This has significant consequences for theoretical advancement (expanding the constructs of assessment), for practical innovation (providing educators with culturally attuned resources), and

for policy implications (reinforcing character-based education efforts in Indonesia). The progress of innovation in Islamic education is set to profoundly impact the successful implementation of collaborative learning methodologies and the evolution of Islamic institutions into dynamic learning organizations (Hidayah, 2025).

D. CONCLUSION

The study offers both theoretical and practical contributions. Theoretically, it advances the construct of speaking assessment by proposing the Religious-Values Infused Speaking Assessment (RV-ISA), which aligns CEFR descriptors with Islamic ethical indicators. This framework broadens assessment beyond conventional linguistic and pragmatic dimensions, responding to calls for more contextually grounded and culturally responsive approaches in language testing. Practically, the findings suggest that teachers require clear and transparent rubrics that explicitly integrate linguistic and value-based criteria, enabling them to provide fair, developmental feedback, and align with the moral mission of Islamic schools.

At the same time, several limitations should be acknowledged. The small sample size, restricted regional scope, and qualitative design limit the generalizability of the findings. Future research should expand to a larger and more diverse population, employ mixed-method approaches, and empirically validate the RV-ISA model to ensure reliability and positive washback effects.

Overall, this study concludes that speaking assessment in Islamic schools is most effective when it balances global communicative standards with local educational values. The RV-ISA provides a promising pathway for advancing language learning and value-based education in faith-based contexts by integrating linguistic accuracy, communicative competence, and ethical character into one framework.

Reference

Books

Brown, H. D., & Abeywickrama, P. (2018). *Language assessment: Principles and classroom practices* (2nd ed.). Pearson.

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.

Fulcher, G. (2010). *Practical language testing*. Routledge. https://doi.org/10.4324/9780203841679

Fulcher, G. (2015). *Re-examining language testing: A philosophical and social inquiry*. Routledge.

Hughes, A. (2003). *Testing for language teachers* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511732980

Luoma, S. (2004). *Assessing speaking*. Cambridge University Press. https://doi.org/10.1017/CBO9780511733017

McNamara, T., Knoch, U., & Fan, J. (2019). *Fairness, justice, and language assessment: The role of measurement*. Oxford University Press.

Journal Articles

Arthur, J. (2008). Traditional approaches to character education in Britain and America. *Journal of Moral Education*, *37*(3), 395–410. https://doi.org/10.1080/03057240802227494

Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy & Practice*, *25*(6), 551–575. https://doi.org/10.1080/0969594X.2018.1441807

Brookhart, S. M. (2018). Appropriate criteria: Key to effective rubrics. *Frontiers in Education*, *3*, Article 22. https://doi.org/10.3389/feduc.2018.00022

Christie, L. S., & Listyani, L. (2018). Teachers' strategies to improve students' self-confidence in speaking: A study at two vocational schools in Central Borneo. *Register Journal*, *11*(2), 121–138. https://doi.org/10.18326/rgt.v11i2.139-153

Cumming, A. (2009). Language assessment in education: Tests, curricula, and teaching. *Annual Review of Applied Linguistics*, *29*, 90–100. https://doi.org/10.1017/S0267190509090084

Dao, T. T., & Nguyen, T. H. T. (2020). Enhancing communication competence via assessment activities of speaking skills from a communicative approach. *LLT Journal: A Journal on Language and Language Teaching*, *23*(2), 293–303. https://doi.org/10.24071/llt.v23i2.2497

Davison, C., & Leung, C. (2009). Current issues in English language teacher-based assessment. *TESOL Quarterly*, *43*(3), 393–415. https://doi.org/10.1002/j.1545-7249.2009.tb00242.x

Denzin, N. K. (2012). Triangulation 2.0. *Journal of Mixed Methods Research*, *6*(2), 80–88. https://doi.org/10.1177/1558689812437186

Duque-Aguilar, A. (2019). Speaking assessment practices in EFL classrooms. *HOW Journal*, *26*(1), 65–84. https://doi.org/10.19183/how.26.1.478

Duque-Aguilar, J. F. (2021). Teachers' assessment approaches regarding EFL students' speaking skills. *Profile: Issues in Teachers' Professional Development*, *23*(1), 161–177. https://doi.org/10.15446/profile.v23n1.85964

Gamlo, N. H. (2019). EFL learners' preferences for corrective feedback in speaking activities. *World Journal of English Language*, *9*(2), 28–37. https://doi.org/10.5430/wjel.v9n2p28

Guo, J. (2024). Collecting language assessment data in the age of pandemic: A preliminary case study of Chinese EFL learners. *Linguistics Vanguard*, *10*(s5), 397–405. https://doi.org/10.1515/lingvan-2021-0148

Halstead, J. M. (2004). An Islamic concept of education. *Comparative Education*, *40*(4), 517–529. https://doi.org/10.1080/0305006042000284510

Hasanah, U., & Hidayah, S. N. (2025). Enhancing Long-term Memory Through Keeping Vocabulary Records Strategy in an Islamic Higher Education:(Meningkatkan Daya Ingat Jangka Panjang melalui Strategi Pencatatan Kosakata di Perguruan Tinggi Islam). *AJMIE: Alhikam Journal of Multidisciplinary Islamic Education*, *6*(1), 104-137.

Hasanah, U., & Nisyak, U. L. D. (2025). Inovasi Edukasi Berbasis Visual: Teknik Picture Sticker Reward dalam Meningkatkan Motivasi Belajar di MA Darunnajah Kelutan Trenggalek. *Al Irfan: Jurnal Ilmu Pendidikan dan Penelitian*, *1*(1), 30-44.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112. https://doi.org/10.3102/003465430298487

Hidayah, S. N. (2025). Universitas sebagai Organisasi Pembelajaran: Studi Kasus di Universitas Islam Negeri Jawa Timur Indonesia. *Al Irfan: Jurnal Ilmu Pendidikan dan Penelitian*, *1*(1), 95-106.

Hidayah, S. N. (2022). Evaluative Analysis On Mooc With an Lms to Improve English Teachers' Teaching Competences. *IJET (Indonesian Journal of English Teaching)*, *11*(1), 30-45.

Hidayah, S. N. (2016). *Mentoring novice secondary teachers as an alternative training program: a case study of a secondary school in Indonesia* (Doctoral dissertation, University of Wollongong).

Hidayah, S. N., & Hasanah, U. (2020). Evaluation of MOOC "Fun Teaching for Fluent Speaking" Implementation to Improve English Teaching Competence. *Southeast Asia Journal on Open and Distance Learning*, *14*(1), 204-204.

Inbar-Lourie, O. (2013). Guest editorial: Language assessment literacy. *Language Testing*, *30*(3), 301–307. https://doi.org/10.1177/0265532213480126

Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity, and educational consequences. *Educational Research Review*, *2*(2), 130–144. https://doi.org/10.1016/j.edurev.2007.05.002

Köroğlu, Z. Ç. (2021). EFL teachers' perspectives on speaking assessment in Turkish secondary schools. *Language Testing in Asia*, *11*(1), 1–17. https://doi.org/10.1186/s40468-021-00126-3

Kusumawati, A., & Sumardi, S. (2020). Teachers' practices in assessing speaking: Challenges and opportunities. *Indonesian Journal of Applied Linguistics*, *10*(2), 381–390. https://doi.org/10.17509/ijal.v10i2.25321

Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research*, *35*(6), 382–385. https://doi.org/10.1097/00006199-198611000-00017

Richards, J. C. (2015). The changing face of language learning: Learning beyond the classroom. *RELC Journal*, *46*(1), 5–22. https://doi.org/10.1177/0033688214561621

Rohmah, G. N., & Hanifiyah, L. (2023). Integrated-holistic teaching prototype for English lessons at Islamic school: A step for building holistic students. *Language Literacy*, *7*(1), 241–254. https://doi.org/10.30743/ll.v7i1.6665

Sakale, S. (2019). The important role of teachers' feedback during speaking activities in Moroccan classes. *Arab World English Journal*, *10*(3), 344–351. https://doi.org/10.24093/awej/vol10no3.24

Sandelowski, M. (2000). Whatever happened to qualitative description? *Research in Nursing & Health*, *23*(4), 334–340. https://doi.org/10.1002/1098-240X(200008)23:4<334::AID-NUR9>3.0.CO;2-G

Scarino, A. (2013). Language assessment literacy as self-awareness: Understanding the role of interpretation in assessment and teacher learning. *Language Testing*, *30*(3), 309–327. https://doi.org/10.1177/0265532213480128

Venera, U. (2017). The design and use of speaking assessment rubrics. *Journal of Education and Practice*, *8*(32), 135–141. https://core.ac.uk/download/pdf/234641197.pdf

Vogt, K., & Tsagari, D. (2014). Assessment literacy of foreign language teachers: Findings of a European study. *Language Assessment Quarterly*, *11*(4), 374–402. https://doi.org/10.1080/15434303.2014.960046

Xu, Y., & Brown, G. T. L. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, *58*, 149–162. https://doi.org/10.1016/j.tate.2016.05.010

Yuzar, E. (2020). Incorporating communicative competence in assessment and English language teaching in multilingual settings. *REILA: Journal of Research and Innovation in Language*, *2*(1), 8–13. https://doi.org/10.31849/reila.v2i1.3864

Webpage with an Author

Muttaqien, S. (2017). The integration of Islamic values in the English language teaching (ELT) in Islamic higher education institution: A case study in STAIN Pamekasan. Universitas Negeri Malang Repository. https://repository.um.ac.id/58846/

Webpage with No Author

Kementerian Agama Republik Indonesia. (2020). Rencana strategis Kementerian Agama tahun 2020–2024. https://ntt.kemenag.go.id/file/file/RENSTRA_KEMENAG_2020-2024.pdf