

AN INVESTIGATION INTO EFL PRIMARY SCHOOL TEACHER'S COMPETENCE IN DESIGNING AND IMPLEMENTING ASSESSMENT FOR YOUNG LEARNERS

Rokhmani¹, Bintang Cahya Dwi Putra^{2*}

^{1,2} Faculty of Teacher Training and Education Universitas Muhammadiyah Prof. DR. HAMKA
bintang.cahya@uhamka.ac.id

Abstract: The implementation of English for Young Learners (EYL) requires teachers to possess not only pedagogical skills but also advanced language assessment literacy (LAL). This descriptive case study investigated primary school English teachers' competence in designing and implementing language assessments for young learners at Sekolah Dasar Islam Ruhama. Data were qualitatively gathered through open-ended electronic questionnaires and artifact documentation from two English teachers with contrasting academic backgrounds and teaching experiences. The findings revealed a noticeable gap between the teachers' high declarative knowledge of assessment concepts and their actual classroom practices. Despite claiming a robust understanding of assessment principles (validity, reliability, and authenticity), traditional structural-based assessments heavily dominated their practices. Grammar and vocabulary were predominantly evaluated through rigid paper-and-pencil tests and rote-learning methods. Furthermore, alternative assessment techniques, such as *self and peer-assessment*, remained completely unprescribed due to a lack of practical guidelines. The study also highlighted that the teacher with a linear postgraduate degree in English education demonstrated more communicative assessment variations than the teacher from a non-linear science background. This study underscores the urgent need for continuous professional development and targeted workshops on child-friendly alternative assessment rubrics to align classroom evaluation with young learners' cognitive development.

Keywords: Alternative Assessment, English for Young Learners, Language Assessment Literacy, Primary School Teachers.

INTRODUCTION

The teaching of English for Young Learners (EYL) at the elementary school level has undergone a significant paradigm shift over the past two decades. In this fast-paced era of globalization, English is no longer viewed merely as an elective subject or a matter of memorizing grammatical structures, but rather as a primary tool for global communication that must be integrated from an early age to build the foundation for 21st-century literacy (Pinter, 2017). The introduction of a foreign language at an early age is based on the critical period hypothesis, which posits that children exhibit high brain plasticity, enabling them to acquire a new language more intuitively and naturally than adult learners (Birdsong, 2018). However, the effectiveness of this early introduction depends entirely on how the learning process is managed in the classroom. The psychological, social, and cognitive characteristics of elementary school-aged children (ages 6–12) are highly unique and dynamic. According to Shin and Crandall (2014), young learners possess abundant curiosity, learn through physical activities, sensory exploration, and direct social interaction. On the other hand, they also have inherent limitations, such as a relatively short attention span, underdeveloped abstract thinking skills, and a tendency to become bored easily with monotonous activities. Therefore, the entire EYL learning ecosystem—from the curriculum, instructional materials, teaching methods, to

the assessment system—must be carefully designed to align with children's developmental stages or be developmentally appropriate (Cameron, 2001; McKay, 2006; Read, 2015).

Among the various pedagogical components within the framework of English for Young Learners (EYL), the aspects of assessment and evaluation occupy a highly crucial position; yet, in practice, they often remain the most neglected, misunderstood, or erroneously synchronized dimensions. Assessment for young learners entails a complex dual role. It is not merely an administrative tool to justify final marks on report cards (summative evaluation), but rather a continuous diagnostic and formative instrument (*assessment for learning*) aimed at monitoring children's linguistic competence in a holistic and humanistic manner (Linse, 2005; Rea-Dickins, 2016). McKay (2006) explicitly asserts that assessments for children must possess high ecological validity, meaning that the evaluation process should be capable of capturing their actual language abilities within a natural context without triggering stress, anxiety, or the fear of failure (*test anxiety*). When children experience elevated anxiety due to rigid testing procedures, their affective filter increases, which in turn impedes their authentic language performance (Krashen, 2013). Consequently, EYL methodologists strongly recommend the utilization of alternative assessment forms. This mode of evaluation encompasses structured observations during classroom activities, portfolios of children's artwork and writing, group project-based assessments, guided self- and peer-assessment, language games (*game-based assessment*), and contextual performance tasks that seamlessly integrate into daily learning activities (Ioannou-Georgiou & Pavlou, 2003; Brown & Abeywickrama, 2019). Through these alternative assessments, teachers can capture a multidimensional portrait of children's language development, spanning receptive skills (listening and reading) and productive skills (speaking and writing) in an enjoyable and non-intimidating manner.

Nevertheless, a profound chasm exists between these ideal theoretical demands and the empirical reality of primary school English language teaching, particularly within the Indonesian educational context. Field studies indicate that many primary school teachers remain entrenched in traditional paradigms that overemphasize the use of conventional, large-scale paper-and-pencil tests—such as multiple-choice worksheets, short answers, and abstract grammar examinations—to measure children's linguistic capabilities (Zein, 2017; Sukmahidayanti, 2021). This traditional assessment approach, centered on low-level cognitive domains, not only fails to record children's authentic communicative competence but also inflicts destructive psychological impacts. The deployment of child-unfriendly tests can stifle intrinsic motivation, diminish self-confidence, and generate early academic trauma toward foreign languages (Copland et al., 2014; Garton et al., 2016). Evaluations that focus excessively on the product or the final outcome disregard the inherently fluctuating nature of children's language acquisition. Affective aspects such as active participation, the willingness to try, and a positive disposition toward the new language—which are the primary catalysts for early childhood language learning success—are frequently eliminated from teachers' assessment schemes because they are deemed difficult to measure quantitatively (Rixon, 2015; Butler, 2019).

Upon closer examination, the root of this misalignment between theory and practice stems from the level of teachers' assessment literacy. A teacher's competence in conducting EYL assessments is not a simple, singular skill, but rather an intricate, multidimensional construct. This competence demands a robust integration of conceptual knowledge regarding the principles of early childhood evaluation, procedural knowledge in designing valid and reliable yet engaging instruments, and practical performance competence to implement them objectively and fairly amidst heterogeneous classroom dynamics (Inbar-Lourie, 2017; Taylor, 2013). A teacher equipped with sound assessment literacy will be capable of achieving constructive alignment among learning objectives, classroom activities, and the assigned assessment tasks (Biggs, 2014). Conversely, a

deficiency in assessment literacy causes teachers to replicate the examination methods they personally experienced in the past, or to blindly adopt exercises from commercial textbooks without prior modification or validation (Vogt & Tsagari, 2014; Sulistiyo et al., 2020).

This dilemmatic condition is further exacerbated by the fluctuating and inconsistent dynamics of English curriculum policies in Indonesia. Since the abolition of English as a compulsory local content subject at the primary school level in the 2013 Curriculum, up to the reinstatement of its status as an elective or local content subject in the *Kurikulum Merdeka*, the position of EYL in Indonesia has perpetually resided in a gray area (Musthafa, 2010; Zein, 2020). This regulatory uncertainty directly impacts the lack of structured, massive, and sustainable professional training provided by the government specifically addressing EYL teaching methodology and assessment for primary school teachers. Consequently, a phenomenon of teacher mismatch occurs in the field. Many primary school English classes are taught by general classroom teachers (graduates of Primary School Teacher Education/PGSD) who lack a background in linguistics and foreign language teaching methodology; alternatively, they are taught by general English Education graduates who are accustomed to the characteristics of adolescent or adult learners (junior/senior high school) and thus apply the same rigid instructional strategies and examination instruments to young children (Hawanti, 2014; Maili, 2018).

Although literature regarding English language teaching in Indonesia is expanding rapidly, academic attention dedicated to the topic of EYL assessment remains highly scarce and fragmented. Thus far, the majority of prior research in Indonesia has leaned heavily toward exploring teaching strategies, utilizing digital media (such as game applications or animated videos) to teach vocabulary, and developing print-based instructional materials (e.g., Cahyati et al., 2019; Lestari, 2021; Rachmajanti, 2022). Research that specifically, deeply, and comprehensively investigates how teachers design and execute assessments at the primary school level is exceptionally rare. The few studies addressing teacher assessment literacy generally select secondary school teachers (junior/senior high school) or university lecturers as subjects, whose curricular ecosystems and learner characteristics are fundamentally different (Mardani & Ashadi, 2022; Priyanti et al., 2023). There is a distinct literature gap regarding the empirical portrait of primary school teachers' practical competence in translating their theoretical understanding of EYL into tangible instrument designs and classroom assessment actions.

To fill this academic void, this study presents a comprehensive and in-depth focus of investigation. This research aims not merely to examine what teachers know on paper, but to dismantle the reality of teacher competence from two interconnected primary dimensions: the ability to design child-friendly assessment instruments, and their effectiveness and challenges in directly implementing those assessments within the classroom. Utilizing a qualitative-descriptive investigative study approach, this research intends to capture this phenomenon holistically, directly from the perspective of practitioners in the field. The outcomes of this study are expected to provide a significant theoretical contribution to the body of knowledge in Indonesian EYL pedagogy, particularly in enriching the literature on authentic assessment for early childhood. Practically, the findings of this research are targeted to serve as a valuable empirical reference for educational policymakers, school principals, and teacher education institutions (LPTK) in designing more contextual, adaptive, and targeted professional teacher training curriculum programs to improve the quality of primary school teachers' assessment literacy in Indonesia.

Operationally, to provide a sharp and focused direction for analysis, this research is designed to answer three fundamental, interrelated questions:

- a. What is the conceptual understanding of primary school English teachers regarding the fundamental principles of assessment and evaluation for young learners?

- b. What is the practical capability of primary school English teachers in designing alternative assessment instruments that are developmentally appropriate?
- c. What are the tangible obstacles and how are the actual practices performed by teachers in implementing these assessment instruments within daily classroom instructional processes?

LITERATURE REVIEW

1.1. Characteristics of Young Learners in Second Language Acquisition

To construct an ecologically valid assessment system, a profound understanding of the fundamental nature of young learners in acquiring a second language is an absolute prerequisite. Within the context of foreign language pedagogy, the term "young learners" generally denotes children between the ages of 6 and 12 who are pursuing their education at the primary school level (Pinter, 2017). Cognitively, drawing upon Jean Piaget's theory of child development, children within this age cohort are transitioning from the pre-operational stage toward the concrete operational stage (Slattery & Willis, 2001). This cognitive positioning implies that they interpret their surrounding environment through the manipulation of tangible objects, direct experiential learning, and physical activities; consequently, they do not yet possess the developmental capacity to internalize abstract concepts or complex, formal grammatical structures (Cameron, 2001).

On the other hand, early childhood language development is profoundly influenced by the socio-constructivist paradigm pioneered by Lev Vygotsky. Vygotsky underscores the significance of the Zone of Proximal Development (ZPD), which conceptualizes the cognitive distance between a child's independent problem-solving capabilities and the potential development achievable under adult guidance or in collaboration with more capable peers (McKay, 2006). Within the context of language acquisition, meaningful interactions and input-rich environments serve as critical scaffolding mechanisms. Children do not acquire a language through the linear memorization of abstract grammatical formulas; rather, they rely on a holistic process wherein they extrapolate global meaning prior to dissecting and understanding isolated linguistic components (Linse, 2005; Shin & Crandall, 2014).

Furthermore, the unique affective characteristics of young learners necessitate careful consideration during the evaluation process. Although children inherently possess intense curiosity and high intrinsic motivation to communicate, they also exhibit a notably brief attention span and are highly susceptible to cognitive fatigue (Read, 2015). Rigid, high-stakes assessment methods can significantly elevate a child's affective filter—a psychological barrier characterized by anxiety and apprehension which, according to Stephen Krashen's theory, impedes the internalization of comprehensible language input within the brain (Krashen, 2013). Consequently, the core principles of EYL pedagogy dictate that language must be presented as a living, engaging tool for communication, rather than an object of theoretical study evaluated through intimidating testing instruments (Copland & Garton, 2014).

1.2. The Nature of Assessment and Evaluation in the Context of EYL

Within pedagogical literature, the terms "assessment" and "evaluation" are frequently utilized interchangeably; nonetheless, they possess fundamentally distinct conceptual boundaries and functional purposes. Brown and Abeywickrama (2019) define assessment as a deliberate, structured, and continuous process of gathering information regarding students' academic performance within the classroom. Conversely, evaluation refers to the decision-making process or the justification of value concerning the overall effectiveness of an educational program or a student's final achievement, formulated based on interpreted assessment data (Rea-Dickins, 2016). At the primary school level, the primary focus of educators should lean heavily toward the domain of assessment, specifically *assessment for learning* (formative assessment), rather than merely *assessment of learning* (summative assessment).

Assessment for young learners embodies a child-friendly, individual growth-oriented philosophy. McKay (2006) formulates several core principles governing EYL assessment:

1. Continuity: Assessment must be seamlessly integrated into the daily instructional process, rather than serving as a detached, high-stakes activity at the end of an academic term.
2. Multidimensionality: Evaluation must capture not only cognitive capabilities but also psychomotor development and the child's affective disposition toward the language.
3. Fairness: Assessment must accommodate diverse student learning styles (visual, auditory, and kinesthetics) and ensure equity for children progressing at varying learning paces.

Traditional assessment, which relies on paper-and-pencil examinations featuring multiple-choice or short-answer formats, is deemed inadequate and developmentally inappropriate for evaluating children's linguistic competence (Butler, 2019). Such rigid examinations merely measure passive *knowledge about language* at a rote-memorization level, rather than a child's capacity for active *language use* in authentic communication. Furthermore, traditional tests have the potential to distort a child's actual potential due to confounding variables, such as limitations in their fine motor writing skills or their inability to comprehend overly complex test instructions (Rixon, 2015).

1.3. Alternative Assessment Paradigm

To mitigate the limitations inherent in traditional testing, EYL experts advocate for the adoption of alternative or authentic assessment paradigms. Alternative assessment is explicitly designed to capture student performance directly through meaningful, contextualized tasks that mirror real-world situations (Ioannou-Georgiou & Pavlou, 2003). The defining characteristic of alternative assessment lies in its balanced emphasis on both process and product evaluation, the active engagement of students in self-reflection, and the provision of constructive, descriptive feedback aimed at fostering subsequent learning improvements (Brown, 2019).

Within the EYL classroom, several forms of alternative assessment have proven empirically effective:

a. Classroom Observation

During classroom observation, educators monitor children's natural interactions during routine activities, such as collaborative group work or language games. Utilizing structured observation checklists, teachers can evaluate listening comprehension, oral fluency, and affective participation without subjecting the children to the psychological pressure of a formal testing environment (McKay, 2006).

b. Portfolio Assessment

A portfolio constitutes a systematic, longitudinal collection of authentic student work, such as labeled illustrations, short poems, audio recordings, or drafting assignments. Portfolios offer robust visual evidence of a child's developmental trajectory over time while simultaneously fostering a sense of ownership and pride in their academic artifacts (Linse, 2005).

c. Task-Based and Project Assessment

In task-based and project assessment, children are assigned performance tasks that require the functional application of language to complete tangible projects. Examples include designing greeting cards, executing basic role-plays, or engaging in narrative storytelling. This evaluative approach integrates multiple macro-language skills concurrently while stimulating the child's creative faculties (Read, 2015).

d. Game-Based Assessment

Given that play is fundamental to childhood development, embedding evaluative components within language games (such as charades, scavenger hunts, or interactive digital quizzes) serves as a strategic mechanism to minimize test anxiety and maximize learners' emotional engagement (Shin & Crandall, 2014).

e. Teachers' Assessment Literacy

The efficacy of implementing these diverse alternative assessment methodologies is ultimately contingent upon the level of assessment literacy possessed by the educator. The term "Assessment Literacy" (AL) was originally popularized by Stiggins (1991) to describe a teacher's comprehension of foundational evaluation principles, the capacity to select and design appropriate assessment methods, and the skill to communicate evaluative outcomes to relevant stakeholders. In subsequent theoretical advancements, Taylor (2013) and Inbar-Lourie (2017) extended the boundaries of AL specifically to language pedagogy—termed Language Assessment Literacy (LAL)—which encompasses knowledge regarding the nature of language, second language acquisition theories, and the socio-political competencies influencing examination policies.

For a primary school English teacher, assessment literacy necessitates an integrative capability known as constructive alignment (Biggs, 2014). Educators must ensure that communicative EYL learning objectives harmonize seamlessly with classroom instructional activities and are evaluated using authentic, applicative instruments. Absent adequate assessment literacy, teachers frequently experience pedagogical dissonance: they deliver instruction using interactive, communicative methods (such as singing and gaming) yet resort to rigid, multiple-choice testing when evaluating student learning outcomes (Vogt & Tsagari, 2014). Consequently, professional teacher development initiatives must look beyond instructional strategies (*how to teach*) and equitably integrate the proficiency required to design and execute assessment instruments (*how to assess*) (Sulistiyo et al., 2020).

1.4. The Policy Landscape and Implementation Challenges of EYL in Indonesia

The practices of EYL pedagogy and assessment in Indonesia are inextricably linked to the volatile dynamics of national curriculum regulations and policies over the past two decades. Historical records indicate that the status of the English language at the primary school level has undergone fluctuations that have introduced significant ambiguity for field practitioners. From the 1994 Curriculum through the 2006 Curriculum (KTSP), English was recognized as a Local Content subject (*Muatan Lokal / Mulok*), routinely instructed beginning from the fourth grade of primary school (Musthafa, 2010; Zein, 2017). However, with the implementation of the 2013 Curriculum, the position of English in primary schools was eliminated from the mandatory curricular framework. This radical policy shift was rationalized by the government's intent to fortify national language literacy (Bahasa Indonesia) and reinforce early childhood character education (Hawanti, 2014). This legislative decision triggered widespread institutional confusion, leading some schools to retain English as an extracurricular activity while others abolished the subject entirely.

A subsequent paradigm shift occurred with the launch of the *Kurikulum Merdeka*. This recent policy framework accommodates the reinstatement of English in primary schools as an elective subject, with prospective plans to transition it into a mandatory course in the near future (Zein, 2020). Although this policy has been widely embraced by the public, prolonged regulatory inconsistencies have left a legacy of deep-seated systemic issues, particularly regarding teacher competence and institutional readiness (Maili, 2018). One of the most palpable impacts of this instability is the widening disparity in instructional quality and the total absence of standardized nationwide assessment frameworks tailored for the primary school level.

Currently, the most formidable challenge confronting Indonesia is the phenomenon of teacher mismatch, characterized by a misalignment in educators' specialized academic backgrounds (Sukmahidayanti, 2021). Out of systemic necessity, numerous primary schools assign general classroom teachers—graduates of Primary School Teacher Education (PGSD)—to instruct English due to the non-existence of dedicated civil service posts for primary school English educators. These classroom teachers frequently demonstrate limited English language proficiency and lack formal training in foreign language teaching methodologies tailored for children (Hawanti, 2014). Conversely, in institutions capable of recruiting dedicated English teachers, the instructional staff predominantly comprises graduates of general English Education departments, whose tertiary curricula are heavily dominated by pedagogical theories geared toward adolescent and adult learners (junior and senior high school levels). Consequently, these educators experience a methodological shock; they exhibit a tendency to duplicate rigid, theoretical-structural instructional techniques and examination instruments and apply them to primary school students who are still operating within the concrete operational stage of cognitive development (Zein, 2020).

Beyond the limitations of formal educational backgrounds, this condition is severely compounded by a scarcity of continuous professional development (CPD) programs specifically addressing EYL assessment literacy, from both regional and central government bodies (Musthafa, 2010). When educators are confronted with the mandates of the *Kurikulum Merdeka*—which prioritizes student-centered, authentic assessments—they lack practical guidelines and concrete exemplars regarding how to construct evaluation rubrics, compile structured portfolios, or execute objective classroom observations. Furthermore, burdensome administrative tasks, high teaching loads, and a lack of supporting media infrastructure in regional schools act as structural impediments. Consequently, teachers frequently resort to pragmatic shortcuts: deploying conventional, traditional written examinations for the sole purpose of simplifying grading and reporting procedures, irrespective of the developmental and psychological mismatch between these methods and their young learners (Sulistiyo et al., 2020; Sukmahidayanti, 2021).

2. Research Methodology

This study employed a qualitative approach utilizing a Descriptive Case Study Design. According to Yin (2018), a descriptive case study design is highly appropriate when the researcher intends to investigate a contemporary phenomenon in-depth directly within its real-life context, particularly when the boundaries between the phenomenon and the context are not clearly evident. Within the scope of this investigation, the researcher conducted an in-depth inquiry to capture the empirical reality of teachers' competencies and constraints in designing and implementing English for Young Learners (EYL) assessments. The primary objective of this research is to delineate the qualitative characteristics of the participants' conceptual understanding, pedagogical instrument products, and practical classroom actions without manipulating or introducing experimental interventions to the research subjects (Creswell & Creswell, 2018).

2.1. Population, Sample, and Data Source

Given that this research employs a qualitative methodology with a case study approach, the concepts of population and sample are interpreted as Contextual Social Units selected through Purposive Sampling. According to Patton (2015), purposive sampling focuses on the strategic selection of information-rich cases that align precisely with the specific criteria and objectives of the study.

The research subjects comprised all active English language teachers instructing at Sekolah Dasar Islam Ruhama. Based on these institutional criteria, a sample of 2 (two) English language teachers was obtained. The primary data source (*Data Source*) for this investigation is digitally recorded and structured within a response questionnaire portfolio document.

The demographic characteristics and academic backgrounds of the two participants were identified as follows:

- a. Participant 1 (Inisial AR): Holds a non-linear undergraduate degree (B.Ed. in Biology Education) with a professional teaching tenure as a primary school English teacher of approximately 1 year.
- b. Participant 2 (Inisial AOS): Holds an undergraduate degree in Economic Accounting and a linear postgraduate Master's degree (M.Pd.) in English Language Education, with a professional teaching tenure as a primary school English teacher of 4 years.

2.2. Research Instrument and Operational Definitions

2.2.1. Research Instrument

In qualitative inquiry, the primary instrument for data collection is the researcher themselves (*human instrument*), who functions as the core agent for gathering empirical evidence (Lincoln & Guba, 1985). Nevertheless, to maintain data objectivity and structural integrity, the researcher was supported by a secondary instrument in the form of a Semi-Structured Electronic Questionnaire integrated with a digital artifact documentation system. This questionnaire framework comprised 34 items evaluating demographic profiles, conceptual understanding of testing dichotomies via structured responses, and qualitative strategies for assessing four macro-language skills alongside core linguistic components (*vocabulary, grammar, and pronunciation*). Additionally, the instrument featured dedicated portals for uploading physical assessment artifacts, specifically mid-term (UTS) and final-term (UAS) examination documents.

2.2.2. Operational Definition

To prevent interpretive ambiguity, the variables within this competency investigation are operationally defined as follows:

- a. **Assessment Design Competence:** The practical proficiency of educators in drafting instructions, test items, word choices, and evaluation rubrics for both mid-term (UTS) and final-term (UAS) examinations that are child-friendly and developmentally appropriate.
- b. **Assessment Implementation Competence:** The empirical strategies deployed by educators to execute macro-language skill evaluations (listening, speaking, reading, and writing) using either conventional testing methods or alternative assessment techniques (such as observations and portfolios) within the primary school classroom.
- c. **Young Learners:** Primary school students within the concrete operational stage of cognitive development (ages 6–12) who receive foreign language instruction at Sekolah Dasar Islam Ruhama.
- d. **Data Collection Procedure**

Data collection in this study was executed systematically across three primary stages to ensure the comprehensiveness of the data corpus (Miles, Huberman, & Saldaña, 2014):

- **Digital Questionnaire Stage:** The researcher administered a digital mapping questionnaire to both participants at Sekolah Dasar Islam Ruhama to harvest declarative data regarding their conceptual understanding of EYL assessment principles.
- **Documentary (Artifact) Study Stage:** The researcher downloaded, compiled, and examined all physical artifacts of the assessment models, including mid-term (UTS) and final-term (UAS) examination exemplars uploaded by the teachers into an integrated Google Drive repository.
- **Contextual Clarification Stage:** The researcher conducted a textual analysis of the written narrative descriptions provided by the teachers concerning their practical methodologies for evaluating students' listening, speaking, vocabulary, pronunciation, grammar, reading, and writing skills in the classroom.

2.2.3. Data Analysis

Qualitative data in the form of narrative texts and examination documents were interactively analyzed utilizing the **Thematic Analysis Model** formulated by Miles, Huberman, and Saldaña (2014). The data analysis process followed a three-stream cyclical pathway outlined below:

1. **Data Condensation:** The researcher sorted, filtered out irrelevant data, simplified, and coded the narrative responses as well as the examination instruments.
2. **Data Display:** The researcher presented the categorized data within descriptive matrices, percentage tables reflecting conceptual understanding, or verbatim narrative excerpts to delineate the comparative competencies between Participant 1 and Participant 2.
3. **Conclusion Drawing and Verification:** The researcher conducted an in-depth interpretation of the emerging patterns, cross-referenced them with EYL assessment literacy theories, and formulated final conclusions to address the primary research questions.

2.2.4. Validity

To ensure the trustworthiness and validity of the qualitative data, this study implemented Data Source Triangulation and Method Triangulation techniques (Denzin, 2017). Data source triangulation was operationalized by contrasting the perceptions and assessment outcomes of two educators possessing contrasting academic backgrounds (Biology Education versus a Master's degree in English Language Education). Concurrently, method triangulation was achieved by cross-referencing self-reported qualitative data (the teachers' written narrative responses) with authentic empirical evidence consisting of the physical mid-term (UTS) and final-term (UAS) examination documents implemented at Sekolah Dasar Islam Ruhama.

2.2.5. Ethical Considerations

This study strictly adheres to the ethical codes of scientific research. Prior to the administration of the research instruments, both participants were provided with a comprehensive explanation regarding the objectives of the study and voluntarily signed an **Informed Consent** form (Bryman, 2016). The researcher provided complete assurances concerning the anonymity of the participants' identities through the use of pseudonyms where necessary within the critical discourse analysis. Furthermore, it was guaranteed that all stored digital portfolio data would be utilized exclusively for the advancement of scientific knowledge and academic journal publication.

FINDINGS AND DISCUSSION

2.3. Discussion

Based on the data, the findings are as follows:

a. Discrepancy Between Declarative Knowledge Claims and Practical Assessment Competence

An intriguing finding of this research illuminates a distinct discrepancy between teachers' declarative knowledge (*knowledge about*) and procedural-practical knowledge (*knowledge how*) within their assessment literacy framework. Conceptually—as evidenced in the closed-ended questionnaire section—both educators affirmatively claimed to comprehend critical assessment principles such as validity, reliability, and authenticity. However, an in-depth exploration of their narrative essay responses regarding the evaluation of linguistic components revealed that traditional-structural assessment models continue to dominate their classroom practices.

For instance, in evaluating grammar and vocabulary, both AR and AOS rely heavily on formal written testing and oral rote-learning methodologies. According to theories advanced by McKay (2006) and Cameron (2001), evaluating the lexical and grammatical competencies of primary school children through rigid, formal written examinations is fundamentally misaligned with their cognitive development, given that they operate within the concrete operational stage. Young learners require context-bound, authentic assessments situated within meaningful scenarios—such as image-matching tasks, role-plays, or the manipulation of tangible objects/realia—rather than the decontextualized memorization of vocabulary lists or the structural reordering of isolated, randomized sentences stripped of functional utility (Linse, 2005).

b. The Impact of Academic Background and Professional Experience on Assessment Design

The divergent assessment performances demonstrated by AR and AOS corroborate the theoretical assertions of Inbar-Lourie (2017) and Vogt and Tsagari (2014) regarding the critical role of Language Assessment Literacy (LAL) acquired through linear professional education pathways. AR, who holds an undergraduate degree in Biology Education and has a limited teaching tenure of one year, exhibits a highly mechanical and textbook-oriented approach to evaluation. This is evidenced by a reliance on static true/false formatting to evaluate listening and reading comprehension. AR's restricted understanding of techniques such as on-the-run assessment suggests that absent a foundational grounding in foreign language pedagogical theories for children (EYL), novice educators encounter substantial difficulties when implementing spontaneous and formative classroom evaluations.

Conversely, AOS, who possesses an advanced academic qualification in the form of a Master's degree in English Language Education alongside four years of teaching experience, demonstrates a considerably richer and more communicative repertoire of evaluative methods. AOS actively explores learners' productive skills by assigning short-story composition tasks (writing) and interactive dialogue practices (speaking).

Nevertheless, AOS's pedagogical practice still reveals certain limitations, specifically the omission of self- and peer-assessment models. This institutional inability or reluctance to engage primary school students in self-reflection or peer evaluation typically stems from an educator's concern regarding grading subjectivity, as well as the inherent complexities of cultivating metacognitive thinking capacities in young learners (Brown & Abeywickrama, 2019; Butler, 2019).

c. Pedagogical Implications for Sustainable Assessment at SD Islam Ruhama

Based on the mid-term (UTS) and final-term (UAS) examination documents submitted by both educators within the response file, there is a clear and pressing need to reconstruct the prevailing paradigm of language testing. The heavy reliance on rigid, large-scale summatively oriented written tests must be progressively minimized. It is highly recommended that institutional authorities, specifically at SD Islam Ruhama, provide targeted interventions through professional workshops focusing on the development of alternative assessment rubrics—such as portfolio evaluation criteria and systematic guidelines for observing young learners' behavioral language use.

Furthermore, educators must be sensitized to the understanding that measuring children's language proficiency does not inherently require rigid, numerically driven examination scoring sheets. Instead, priority should be given to fostering a constructive and descriptive washback effect. Providing such formative, qualitative feedback ensures that young learners' intrinsic motivation to acquire a foreign language from an early stage is not undermined by an apprehension toward formal evaluation (Read, 2015; Sulistiyo et al., 2020).

2.4. Result

2.4.1. Teachers' Conceptual Understanding of EYL Assessment Principles Guru Mengenai Prinsip Asesmen EYL

Based on the primary data obtained, the conceptual understanding of both educators at Sekolah Dasar Islam Ruhama demonstrates a uniform claim of high-level comprehension regarding the foundational aspects of evaluation. Both Participant 1 (AR) and Participant 2 (AOS) affirmatively stated that they comprehend the concepts of evaluation, assessment, testing, the principles of effective assessment, and the various purposes of assessment. Furthermore, they confirmed their

understanding of the formative-summative assessment dichotomy, as well as the informative, diagnostic, and evaluative functions of language testing.

However, a divergence in conceptual understanding began to manifest when queries were directed toward technical terminology and alternative assessment techniques:

- Principles of Language Assessment:** Both educators claimed to understand the core tenets of *practicality, reliability, validity, authenticity, and washback*.
- Alternative Assessment:** Regarding specific evaluation techniques, AR (the educator with a background in Biology Education) acknowledged a lack of comprehension concerning the concept of *on-the-run assessment* (spontaneous, real-time classroom evaluation). Conversely, AOS (the graduate with a Master's degree in English Language Education) confirmed a comprehensive understanding of all proposed assessment methodologies.

2.4.2. Praktik The Implementation Practices of Alternative Assessment Techniques in Primary Schools

Although both participants claimed a high level of comprehension at the theoretical level, their actual implementation track records within the classroom reveal a significant discrepancy. The table below summarizes the comparison between the participants' conceptual understanding and their self-reported classroom practices across a variety of child-friendly assessment techniques for young learners:

Types of Alternative Assessment Techniques	Participant 1 (AR) (Biology Education / 1 Year of Teaching Experience)	Participant 2 (AOS) (Master's in English Language Education / 4 Years of Teaching Experience)
Observation Assessment	Comprehended & Implemented	Comprehended & Implemented
On-the-run Assessment	Not Comprehended & Not Implemented	Comprehended & Implemented
Conferences Assessment	Comprehended & Not Implemented	Comprehended & Implemented
Portfolio Assessment	Comprehended & Implemented	Comprehended & Implemented
Self & Peer Assessment	Comprehended & Implemented	Comprehended & Implemented

Based on the aforementioned table, it is evident that Self and Peer Assessment constitutes the most unfamiliar technique to be operationalized, with both educators uniformly stating that they have never implemented this methodology within the classrooms of SD Islam Ruhama.

2.4.3. Pedagogical Approaches to Evaluating Learners' Language Skills and Linguistic Components

The narrative descriptions regarding the operationalization of daily formative assessments across the four macro language skills and three linguistic components reveal pronounced procedural variations. These discrepancies appear to be heavily influenced by the distinct academic backgrounds of each educator:

- a. Listening Comprehension: AR predominantly relies on audio playback followed by the administration of rigid evaluation instruments in the form of true/false statements. Conversely, AOS integrates audio playback with live, teacher-led vocabulary pronunciation using an amplification system.
- b. Speaking Proficiency: AR constrains oral evaluation strictly to text-reading tasks to assess pronunciation, fluency, and intonation. AOS adopts a more interactive trajectory, assessing students individually through story narration or by facilitating paired interactive dialogue performances at the front of the classroom, adhering to textbook exemplars.
- c. Vocabulary Acquisition: AR relies exclusively on formal, written diagnostic instruments. In stark contrast, AOS employs a structured, traditional rote-learning approach, wherein students are assigned incremental sets of vocabulary to memorize at home, which they must then systematically recite orally during the subsequent class session.
- d. Pronunciation: AR evaluates this component via oral text-reading examinations. AOS requires students to read passages or dialogues aloud, while simultaneously prompting them to spontaneously construct novel sentences utilizing a provided set of vocabulary words.
- e. Grammatical Competence: Both participants uniformly adopt a conventional structural approach, utilizing formal written examinations or structural exercises extracted from student workbooks.
- f. Reading Comprehension: AR implements a text-comprehension model coupled with true/false verification items. AOS applies text- and dialogue-reading assignments followed by context-bound, independent sentence construction tasks.
- g. Writing Skills: AR evaluates learners by presenting jumbled sentences to be reordered into complete, structurally sound simple sentences. In contrast, AOS implements a more expressive writing task, requiring the composition of simple, original short stories aligned with the thematic content of the lesson.

CONCLUSION AND SUGGESTION

2.5. Conclusions

Based on an in-depth investigation into the competencies of English language educators at Sekolah Dasar Islam Ruhama, this study draws several foundational conclusions. First, at the theoretical conceptual level, the educators demonstrate a uniform claim of high-level comprehension regarding the dichotomies of assessment, testing, and evaluation, as well as the formative-summative functions of language testing. However, proficiency in specific technical terminology and alternative

assessment techniques remains linearly influenced by their specialized educational background in teaching English to young learners (EYL).

Second, within the dimensions of assessment design and implementation, a pronounced discrepancy exists between theoretical claims and practical classroom execution. Although the educators comprehend the core essence of daily formative assessment, daily evaluative instruments and formal examinations—such as mid-term (UTS) and final-term (UAS) tests—are still heavily dominated by rigid, traditional-structural approaches. These include written worksheets, conventional grammar exercises, and mechanical vocabulary memorization lists, which are fundamentally misaligned with the concrete operational stage of primary school students' cognitive development.

Finally, student-centered, autonomy-based evaluation techniques, such as self- and peer-assessment, constitute the most unfamiliar domain and remain entirely un-implemented by the educators in the field, primarily due to a lack of tactical guidelines regarding alternative assessment criteria.

2.6. Suggestions

To bridge the pedagogical discrepancies identified within this study, several strategic recommendations are proposed to the relevant stakeholders:

- a. For Primary School English Educators: It is highly recommended that teachers begin shifting their evaluative paradigm from an *assessment of learning* framework (which heavily emphasizes rigid, product-oriented numerical scoring) toward a child-friendly *assessment for learning* approach. Educators should imperceptibly integrate assessment elements through authentic performance tasks. These include activity-based group observations, portfolios of creative work, and simple communicative projects designed to mitigate psychological test anxiety among young learners.
- b. For School Administrators and Management (SD Islam Ruhama): Institutional authorities are expected to facilitate Continuous Professional Development (CPD) programs specifically tailored to language assessment literacy for young learners. There is an urgent need to conduct periodic professional workshops addressing the design of alternative assessment rubrics and the management of digital portfolios, with a particular focus on supporting educators who lack a linear academic background in EYL education.
- c. For Future Researchers: Given that this study utilized a descriptive case study design with a constrained sample size of two educators, future researchers are encouraged to expand the scope of investigation. This can be achieved by employing macro-scale survey methods or utilizing action research (PTK) designs to empirically test the efficacy of implementing specific alternative assessment models within primary school classrooms.

REFERENCES

- Biggs, J. (2014). Constructive alignment in university teaching. *HERDSA Review of Higher Education*, 1(1), 5-22.
- Birdsong, D. (2018). Plasticity, age, and the critical period hypothesis. *Wiley Blackwell Handbook of Bilingualism*, 143-159.
- Brown, H. D., & Abeywickrama, P. (2019). *Language assessment: Principles and classroom practices* (3rd ed.). Pearson Education.

- Bryman, A. (2016). *Social research methods* (5th ed.). Oxford University Press.
- Butler, Y. G. (2019). Assessment of young learners. *The Routledge Handbook of Teaching English to Young Learners*, 339-355.
- Cahyati, S. S., Parmawati, A., & Atmawidjaja, N. S. (2019). Optimizing English teaching and learning process to young learners. *Journal of Educational Experts (JEE)*, 2(2), 107-114.
- Cameron, L. (2001). *Teaching languages to young learners*. Cambridge University Press.
- Copland, F., & Garton, S. (2014). Key themes and future directions in teaching English to young learners. *TESOL Quarterly*, 48(4), 653-662.
- Copland, F., Garton, S., & Burns, A. (2014). Challenges in teaching English to young learners: Global perspectives and local realities. *TESOL Quarterly*, 48(4), 738-762.
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.
- Denzin, N. K. (2017). *The research act: A theoretical introduction to sociological methods*. Routledge.
- Garton, S., Copland, F., & Burns, A. (2016). Investigating assessment in meaningful contexts for young learners. *British Council English Language Teaching Research Papers*, 16(2), 1-24.
- Hawanti, S. (2014). Implementing English in Indonesian primary schools: Teachers' viewpoints. *International Journal of Primary Education*, 13(1), 51-64.
- Inbar-Lourie, O. (2017). Language assessment literacy. In *Language testing and assessment* (pp. 257-270). Springer, Cham.
- Ioannou-Georgiou, S., & Pavlou, P. (2003). *Assessing young learners*. Oxford University Press.
- Krashen, S. (2013). The affective filter hypothesis and its injury to language acquisition. *Language Learning Journal*, 41(2), 112-121.
- Lestari, S. (2021). Developing English supplementary reading materials for young learners. *Journal of English Language Teaching*, 10(2), 85-94.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. SAGE Publications.
- Linse, C. T. (2005). *Practical English language teaching: Young learners*. McGraw-Hill.
- Maili, S. (2018). Bahasa Inggris pada sekolah dasar: Sebuah tinjauan kompetensi guru. *Jurnal Pendidikan Bahasa*, 7(1), 33-46.
- Mardani, R., & Ashadi, A. (2022). Investigating EFL teachers' assessment literacy and their classroom practices. *Jurnal Pendidikan Bahasa dan Sastra*, 22(1), 45-58.
- McKay, P. (2006). *Assessing young language learners*. Cambridge University Press.
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). SAGE Publications.
- Musthafa, B. (2010). Teaching English to young learners in Indonesia: Essential requirements. *Educationist*, 4(2), 120-125.
- Patton, M. Q. (2015). *Qualitative research & evaluation methods: Integrating theory and practice* (4th ed.). SAGE Publications.
- Pinter, A. (2017). *Teaching young language learners* (2nd ed.). Oxford University Press.
- Priyanti, N. K., Artini, L. P., & Padmadewi, N. N. (2023). Assessment literacy of EFL teachers in secondary schools. *Journal of Language Education and Research*, 9(2), 112-126.
- Rachmajanti, S. (2022). Material development for young learners: Challenges and ways out. *Indonesian Journal of ELT*, 14(1), 55-68.
- Read, C. (2015). *Seven things to know about teaching young learners*. Macmillan Education.
- Rea-Dickins, P. (2016). Assessment in early language learning contexts. *The Handbook of Second Language Assessment*, 229-246.

- Rixon, S. (2015). Primary English policy and implementation in Europe. In *International perspectives on teaching English to young learners* (pp. 31-48). Palgrave Macmillan, London.
- Shin, J. K., & Crandall, J. A. (2014). *Teaching young learners English: From theory to practice*. National Geographic Learning/Cengage Learning.
- Slattery, M., & Willis, J. (2001). *English for primary teachers: A handbook with activities and classroom language*. Oxford University Press.
- Stiggins, R. J. (1991). Assessment literacy. *Phi Delta Kappan*, 72(7), 534-539.
- Sukmahidayanti, T. (2021). The utilization of paper-and-pencil test in primary English classrooms. *Journal of English and Education*, 9(3), 201-215.
- Sulistiyo, U., Aliyyah, R. R., & Warni, S. (2020). Evaluating the competence of primary school English teachers in designing assessment. *Journal of English Education and Linguistics Studies*, 7(2), 211-235.
- Taylor, L. (2013). Communicating the theory and practice of language assessment literacy. *Language Testing*, 30(3), 403-412.
- Vogt, K., & Tsagari, D. (2014). Assessment literacy of foreign language teachers: Findings from a European study. *Language Assessment Quarterly*, 11(4), 374-402.
- Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). SAGE Publications.
- Zein, M. S. (2017). Elementary English education in Indonesia: Policy and implementation. *Language Policy*, 16(4), 463-479.
- Zein, R. (2020). *English education for young learners in Indonesia: A focus on teacher professional development*. Routledge.