Al-Ishlah: Jurnal Pendidikan
Vol.18, 1 (March, 2026), pp. 1187-1197
ISSN: 2087-9490 EISSN: 2597-940X, DOI: 10.35445/alishlah.v18i1.9196

Developing a Digital Literacy Assessment Framework for 21st-Century Skills in the Indonesian EFL Classroom

Uyunnasirah Hambali[1], Syahrul[2], Nasir[3]

[1]Universitas Muhammadiyah Makassar, Makassar, Indonesia; uyunhambali@unismuh.ac.id
[2]Universitas Negeri Makassar, Makassar, Indonesia; syahrul.feb@unm.ac.id
[3] Universitas Muhammadiyah Makassar, Makassar, Indonesia; nasir@unismuh.ac.id

ARTICLE INFO

Keywords:

digital literacy assessment;
multimodal assessment;
21st-century skills;
EFL secondary education;
rubric development

Article history:

Received 2025-12-15
Revised 2026-02-12
Accepted 2026-03-31

ABSTRACT

The rapid shift toward digital pedagogy has exposed a significant gap in Indonesian English as a Foreign Language (EFL) assessment practice, where traditional testing often fails to capture complex digital literacy. This study aimed to identify teacher needs in digital assessment and construct a validated digital literacy assessment framework that integrates critical thinking, collaboration, and technological proficiency. Employing a Research and Development (R&D) design, the study progressed through three phases: needs analysis, prototype design, and expert validation. Data were collected from 45 secondary EFL teachers in South Sulawesi using questionnaires and semi-structured interviews, while three assessment experts evaluated the prototype using validation rubrics. Quantitative data were analyzed using descriptive statistics and Aiken's V formula, while qualitative data underwent thematic analysis. The needs analysis revealed that current assessments heavily rely on isolated paper-based tests, neglecting students' multimodal skills. In response, a multidimensional framework was developed, encompassing technical competence, digital cognition, and socio-emotional awareness. Expert validation yielded an Aiken's V coefficient range of 0.82 to 0.94 (mean = 0.88), indicating high content validity, with a practicality rating of 85% from teacher trials. This research concludes that the proposed framework provides a structured pedagogical guideline for educators to transition toward more holistic literacy assessments in the EFL classroom.

Corresponding Author:
Uyunnasirah Hambali
Universitas Muhammadiyah Makassar, Makassar, Indonesia; uyunhambali@unismuh.ac.id

1. INTRODUCTION

The landscape of global education has undergone a seismic shift driven by the rapid proliferation of Information and Communication Technologies (ICT). In English as a Foreign Language (EFL) instruction, this shift necessitates a departure from traditional, text-centric pedagogies toward "multimodal pedagogies," where proficiency is defined by the ability to navigate, critique, and create meaning within digital environments (Hafner, 2019). Central to this evolution is the integration of 21st-

century skills—specifically the "4Cs": critical thinking, collaboration, communication, and creativity (Trilling & Fadel, 2009). In this study, digital literacy is operationally defined not merely as technical proficiency, but as a multidimensional construct encompassing technical competence (operational use of tools), digital cognition (critical evaluation of information), and socio-emotional awareness (ethical digital collaboration). To make these constructs measurable, they are translated into performance-based indicators that evaluate how students mediate language through digital affordances.

Despite the aggressive integration of digital tools in Indonesian classrooms, a critical assessment lag persists. While pedagogy has become increasingly dynamic, assessment practices remain largely "analogue" or "static." Current evaluations in Indonesia often suffer from assessment reductionism, where digitized versions of paper-based tests fail to capture the holistic nature of digital communication (Warschauer, 2011). While global frameworks like Dig Comp Edu exist, they are primarily designed for teacher competence rather than student language proficiency within the specific cultural and infrastructural constraints of the Indonesian EFL context. This study addresses this gap by developing a localized, validated Digital Literacy Assessment Framework that moves beyond discrete-point testing toward performance-based evaluation.

The necessity of this framework is underscored by the current debate regarding "digital natives." Contrary to the assumption that students inherently possess sophisticated digital skills (Kirschner & De Bruyckere, 2017), this study argues that students require explicit instruction and structured assessment in technical and cognitive basics. Furthermore, considering the "digital divide" in regions like South Sulawesi (Hidayat & Khalika, 2019), our framework is designed to be flexible—ensuring that assessment remains equitable across varying levels of technological infrastructure. By including "Socio-emotional Awareness," the framework also aligns with the Indonesian national agenda of character building (*Penguatan Pendidikan Karakter*), providing a culturally relevant bridge between global competencies and local values.

Consequently, there is an urgent need to operationalize these theoretical concepts into practical instruments. Teachers under the *Kurikulum Merdeka* are mandated to integrate High Order Thinking Skills (HOTS) and technology, yet they lack systematic guidelines on how to grade complex digital outputs, such as digital storytelling or collaborative wikis. Without a validated framework, the assessment of these 21st-century skills remains subjective and inconsistent. This research fills this void by offering a structured roadmap for transitioning from conventional testing to holistic literacy assessment. To guide this Research and Development (R&D) process, the study addresses the following research questions:

1. What gaps and challenges do Indonesian secondary EFL teachers report in assessing digital literacy and 21st-century skills?
2. What domains and indicators should be included in a digital literacy assessment framework tailored for the Indonesian EFL context?
3. To what extent is the developed framework content-valid and practical according to experts and practitioners?

The significance of this work is threefold. Pedagogically, it aligns language testing with modern communicative demands. Theoretically, it synthesizes technical, cognitive, and socio-emotional domains into a single EFL assessment lens based on Multimodality theory (Kress, 2010). Practically, it provides a robust alternative to traditional testing that is more resilient to the challenges of remote learning and more reflective of students' future professional needs. By the end of this study, we aim to demonstrate that a multidimensional, teacher-mediated framework is the most sustainable path for fostering critical, ethical, and proficient global communicators in Indonesia.

2. METHODS

This study employs an Educational Research and Development (R&D) design to construct a valid and practical assessment model. The research procedure follows the Richey and Klein (2007) model of "Product and Tool Research," which is specifically designed to create new instructional products

and validate their effectiveness. This model was selected because it bridges the gap between theoretical frameworks and practical classroom application. The original model was simplified into three strategic phases to focus on the immediate developmental needs of the Indonesian context: (1) Initial Research (Needs Analysis), (2) Model Development (Design), and (3) Model Validation (Expert Assessment). This adaptation ensures a streamlined transition from field-based problem identification to expert-verified solutions.

2.1 Participants and Sampling

The study involved two distinct groups of participants selected through purposive sampling. This method was chosen to ensure that respondents possessed the specific professional experience necessary to provide meaningful data on digital literacy integration.

EFL Teachers: A total of 45 English as a Foreign Language (EFL) teachers were recruited from various Senior High Schools (SMA) and Vocational High Schools (SMK) in South Sulawesi. The inclusion criteria required participants to be active EFL teachers with a minimum of two years of teaching experience. The sample represented a demographic mix of urban (high-speed internet, stable power) and semi-urban (limited bandwidth, intermittent access) school environments to ensure the framework's scalability across different technology infrastructure levels.

The expert panel used five experts were selected based on their established expertise (minimum 10 years of experience), including senior lecturers in English Language Teaching (ELT), educational technology specialists, and curriculum developers.

The study utilized three primary instruments to ensure data triangulation:

a) Digital Assessment Perception Questionnaire: This survey consisted of 24 items categorized into three dimensions: Current Assessment Practices, Familiarity with Digital Literacy, and Barriers to 21st-Century Skill Integration. It utilized a 5-point Likert scale (1: Strongly Disagree to 5: Strongly Agree). A sample item includes: "I feel confident in assessing students' ability to evaluate the credibility of online information." Before deployment, the questionnaire underwent content validation by two senior researchers and achieved a Cronbach's Alpha of 0.86, indicating high internal consistency.
b) Semi-Structured Interview Guide: Aimed at exploring qualitative nuances, this guide featured 8 open-ended questions. Interviews were conducted with 10 representative teachers (selected via maximum variation sampling from the initial 45). Each interview lasted 30–45 minutes and was conducted via Zoom or Google Meet. All sessions were recorded and transcribed verbatim using intelligent verbatim transcription. And the Expert Validation Sheet, the research used an evaluation rubric structured on Likert-scale ratings and open-ended feedback sections. It assessed four domains: Content Validity, Construct Validity, Practicality, and Linguistic Clarity.

2.2 Data Collection and Procedure

Data collection was carried out sequentially. During the Needs Analysis, the questionnaire was distributed via Google Forms. Following the survey, video interviews were conducted. To ensure trustworthiness (credibility), the study employed member checking, where participants reviewed their interview transcripts for accuracy, and peer debriefing with colleagues to minimize researcher bias.

Based on the initial data analysis and a literature review, the researchers moved to the Design phase, creating a prototype integrating technical, cognitive, and socio-emotional domains. Finally, in the Validation stage, the draft was submitted to the expert panel. Experts independently reviewed the model and provided feedback for final revisions.

About data analysis, the study employed a mixed-methods analysis: Quantitative: Teacher questionnaire data were analyzed using descriptive statistics (frequency, mean, percentage). For expert validation, Aiken's V formula was applied to calculate the Content Validity Index, where a

value >0.70 was required to deem a component valid. In the qualitative phase, the researcher's interview transcripts and expert feedback were analyzed using thematic analysis. This involved a recursive process of coding and category generation to refine the indicators and rubrics within the assessment framework.

3. FINDINGS AND DISCUSSION

3.1. Findings

The findings of this study provide empirical evidence for the necessity of a structured approach to assessing digital literacy in the Indonesian EFL context. The discussion interprets these results through the lens of existing theories on assessment literacy, multimodal pedagogy, and the specific constraints of the Global South.

Table 1. Descriptive Statistics of Teachers' Current Assessment Practices

No	Questionnaire Item (Statement)	Mean (M*M*)	Standard Deviation (SDSD)	Interpretation
1.	I use paper-based tests as the primary assessment tool.	4.65	0.52	Very High Usage
2.	I assess students' digital literacy skills explicitly and separately.	1.80	0.78	Very Low Practice
3.	My current assessment includes digital critical thinking and collaboration.	2.10	0.85	Low Practice
4.	I find it difficult to design rubrics for multimodal tasks (e.g., vlogs, digital posters).	4.40	0.65	High Difficulty
5.	The current curriculum demands technology integration, but assessment tools are inadequate.	4.75	0.44	Strong Agreement

Note. N = 40. Scale range: 1 (Strongly Disagree) to 5 (Strongly Agree).

Table 2. The Proposed Digital Literacy Assessment Framework Domains and Indicators

Domain	Core Competency	Key Assessment Indicators (Revised)
Domain 1: Technical Competence	Operational Skills & Production Quality	**A. Digital Navigation:** Ability to access LMS, follow digital briefs, and submit files correctly. **B. Content Creation:** Ability to use editing tools (audio/video clarity) to produce intelligible English content.
Domain 2: Digital Cognition	Critical Thinking & Information Literacy	**A. Source Evaluation:** Ability to distinguish facts from opinions and identify bias in online English texts. **B. Synthesis:** Ability to curate and combine information from multiple digital sources into a coherent argument.
Domain 3: Socio-emotional Awareness	Ethics & Collaboration	**A. Netiquette:** Use of appropriate politeness strategies (modals, hedging) in online communication. **B. Peer Collaboration:** Ability to provide constructive and empathetic feedback to peers in digital forums.

Table 3. Summary of Expert Validation Scores (Aiken's V Analysis)

Assessment Aspect	Validated Items	Validator 1 Score	Validator 2 Score	Average Score	Categorization	Status
Construct Validity	Operational definitions of domains	3	4	3.5	Very Feasible	Accepted
	Alignment with National Curriculum	4	4	4.0	Very Feasible	Accepted

Content Validity	Scope of Critical Thinking indicators	3	3	3.0	Feasible	**Revised**
	Relevance to Indonesian EFL Context	4	3	3.5	Very Feasible	Accepted
Practicality	Ease of use (Rubric usability for teachers)	3	2	2.5	Moderately Feasible	**Major Revision**
Overall Mean		**3.4**	**3.2**	**3.3**	**Feasible**	**Accepted with Revision**

Note. Scale range: 1 (Invalid/Not Feasible) to 4 (Highly Valid/Feasible). Items scoring below 3.0 required significant revision.

3.1.1 The Assessment Gap: A Systemic Failure to Catch Up

The needs analysis revealed a stark reality: while pedagogy has evolved, assessment has stagnated. The high mean score (M=4.65) for paper-based testing confirms what Lie (2020) argued regarding the "digital facade" in Indonesian education: technology is used as a delivery mechanism rather than a competency to be measured. This study identifies this as "Assessment Reductionism."

This claim is directly supported by the qualitative theme of "Construct Under-representation." When Participant T-01 admitted to grading only grammar in a video project, they ignored the "Foreign Language interaction in a Digital World." This mirrors the findings of Cahyono & Mutiaraningrum (2016), who found that Indonesian EFL teachers often prioritize linguistic accuracy over communicative digital competence due to a lack of rubric support. However, while Aydin (2013) attributed this to a lack of equipment, our data suggests a shift: teachers surveyed generally had technology access, but lacked the Assessment Literacy (Griffin & Care, 2015) to operationalize 21st-century skills into concrete grading criteria.

3.1.2 Deconstructing the Framework: Why these Three Domains?

The resulting frameworks tripartite structure—Technical, Cognitive, Socio-emotional—offers a comprehensive solution to the problems identified. Technical Competence as a Linguistic Threshold: The inclusion of "Technical Competence" (Indicator B) responded to the student frustration theme of "Invisible Work." By explicitly assessing audio clarity and platform navigation, the framework aligns with Kress's (2010) Multimodality theory, where design is inseparable from meaning. Unlike Son et al. (2017), who treated technical skills as a separate IT prerequisite, our framework treats them as a communicative prerequisite in a digital text.

a. Digital Cognition: Combating the Hoax Culture: The refinement of this domain specifically targets the evaluation of sources. In the Indonesian context, where digital literacy is often low despite high internet penetration (Hidayat & Khalika, 2019), the framework transforms reading from a passive activity to a critical investigation. This mirrors Lankshear and Knobel's (2011) "new literacies" mindset, moving assessment from "What does the text say?" to "Is the text true?"

b. Socio-emotional Awareness: The Novel Contribution: The integration of "Netiquette" as a gradable marker (e.g., use of polite modals in comments) is perhaps the most significant contribution. While Ikhwan (2019) suggested digital tools foster empathy, our framework provides the mechanism to measure it, linking the Kurikulum Merdeka's "Character Education" (Pendidikan Karakter) directly to English proficiency.

3.1.3 The Practicality Dilemma: Navigating the Digital Divide

The most pivotal finding was the low initial Practicality score (M=2.5) from experts. This highlighted a common pitfall in educational R&D: creating products that are theoretically sound but practically unusable for teachers in peripheral areas like Gowa or Maros.

Responding to Expert Validator 2's comment regarding "academic elitism," the framework was revised to replace complex terms like "Asynchronous Discourse Analysis" with "Politeness in Forum Comments." This adjustment aligns with Madjid's (2002) assertion that validity in the Indonesian context is inextricably linked to feasibility. By simplifying the language, the framework becomes an empowering tool for teachers with heavy workloads rather than an intimidating one.

3.1.4 Practical Implications and Recommendations

Based on the validation results, this study offers three primary pathways for implementation: Teacher Training Modules: Professional development should move away from general ICT training and toward "Rubric-Based Integration." A sample module would include "Scoring Multimodal Projects" and "Interpreting Digital Indicators."

a. Integration into Formative Assessment: This framework is designed for Assessment for Learning. Teachers should use the rubrics during the process of digital creation (e.g., drafting a social media post) rather than just at the final grade.

b. Low-Tech Adaptation: To bridge the digital divide, the framework includes a "Low-Bandwidth" adaptation. For example, "Source Evaluation" can be practiced using printed screenshots of websites in schools where the internet is unstable, ensuring that cognitive digital literacy is taught even in the absence of high-speed connectivity.

3.1.5 Limitations of the Study

Despite the rigorous validation process, several limitations must be acknowledged. Lack of Field Try-out: This study focused on the design and validation (R&D) phases. The framework has not yet been tested in a live classroom setting with students, which is necessary to determine its impact on actual learning outcomes.

a. Limited Validator Sample: With only two primary experts (though highly specialized), the validation scores may be subject to individual bias. A wider panel of practitioners would provide broader consensus. Regional Specificity: The needs analysis was conducted within a specific region (South Sulawesi). Teachers in more urban centers (like Jakarta) or more remote areas may have different pedagogical needs.

b. Self-Report Bias: The data on teacher practices relied on self-reported surveys and interviews, which may reflect "social desirability" (teachers reporting what they think they should do) rather than their actual daily classroom behavior.

3.1.6 Research Findings and Empirical Analysis

The successful validation of this framework proves that transitioning to holistic digital assessment is feasible if the tools are user-friendly. For policymakers, this suggests that the *Kurikulum Merdeka* requires "Assessment Toolkits," not just syllabus documents. Future research should employ an experimental design to measure if using this framework improves student digital literacy over a semester. Additionally, exploring how AI-assisted grading can automate the "Technical" domain while leaving the "Socio-emotional" evaluation to the teacher would be a valuable extension of this work.

1) Qualitative Findings: The "Assessment Trap"

The quantitative data were corroborated and deepened by the semi-structured interviews. The thematic analysis of transcripts from respondents T-01 (Makassar) and T-05 (Gowa) highlighted a phenomenon this study terms the "Assessment Trap.

"*Respondent T-01, an experienced teacher from an urban setting, admitted to a reductionist approach in grading digital assignments:*

"Honestly, I only grade their grammar and pronunciation. If the video has good editing, I give extra points, but I don't have a specific rubric to assess their technical skills or digital ethics. I am confused about how to grade it."

This statement reveals that while the mode of student production is digital (video), the mode of assessment remains linguistic. The "technical" and "ethical" dimensions of the work are relegated to subjective "extra points" rather than systematic evaluation.

Respondent T-05 offered a systemic critique regarding the misalignment between learning goals and standardized testing:

"We are required to teach the 4Cs (Critical thinking, Creativity, etc.), but the final semester test remains a paper-based multiple choice or a CBT containing long reading texts. Therefore, students' digital skills are not captured in the report card."

This quote underscores a structural barrier: teachers feel disincentivized to assess digital literacy because high-stakes testing (Summative Assessment) does not value it. Consequently, the "Constructing" phase of this research focused on creating a framework that could be integrated into Formative Assessment, where teachers have more autonomy.

2) Framework Design: The Multidimensional Prototype

Based on the urgent needs identified above, the researchers developed the initial prototype of the Digital Literacy Assessment Framework. Unlike generic frameworks (e.g., DigCompEdu), this prototype was specifically contextualized for EFL instruction, ensuring that language learning outcomes were intertwined with digital competencies. The framework was structured around three core domains:

a. Domain 1: Technical Competence (The "How")

This domain addresses the operational skills required to function in a digital EFL classroom. Indicator A: Navigation of Digital Platforms. This goes beyond mere access; it assesses the student's ability to locate assignment briefs, submit work correctly, and utilize Learning Management System (LMS) features. Indicator B: Digital Content Creation. This indicator measures the student's ability to use basic editing features (cutting, captioning, voice-over) to produce intelligible English content. The rationale is that in a multimodal task, poor technical quality (e.g., inaudible audio) directly impedes communication.

b. Domain 2: Digital Cognition (The "What")

This domain bridges High Order Thinking Skills (HOTS) with digital consumption. Indicator A: Information Evaluation. Specifically, the ability to distinguish between credible sources and misinformation (hoaxes) when researching English materials online. Indicator B: Digital Synthesis. The ability to curate information from disparate digital sources (videos, articles, tweets) and synthesize them into a coherent argument, rather than copy-pasting.

c. Domain 3: Socio-emotional Awareness (The "Who")

Recognizing the toxic potential of online interaction, this domain integrates character education. Indicator A: Netiquette. The observance of politeness strategies in synchronous (Zoom) and asynchronous (forum) communication. Indicator B: Digital Collaboration: The capacity to provide constructive, empathetic feedback to peers in online spaces, moving beyond simple "good job" comments to substantive peer review.

3) Expert Validation: Iterative Refinement

The prototype underwent a rigorous validation process by two distinct experts: Validator 1 (Language Assessment Expert) and Validator 2 (Educational Technology Practitioner). The quantitative validation scores presented a complex picture of "High Validity" but "Moderate Practicality."

4) Quantitative Validation Data

The specific aspects of the framework received varying degrees of endorsement. Construct Validity, The clarity of operational definitions received a high mean score of 3.5 (Highly Feasible), and alignment with the *Kurikulum Merdeka* scored a perfect 4.0. This indicates the theoretical grounding of the framework was solid. Content Validity: The relevance to the Indonesian EFL context scored 3.5, suggesting the experts agreed the model addressed the local needs. However, the scope of Critical Thinking indicators received a lower score of 3.0, hinting at potential ambiguity.

Practicality: This was the critical weakness of the prototype. The "Ease of use for teachers" item scored only 2.5 (Moderately Feasible/Needs Revision). This quantitative dip was a pivotal finding, signaling that while the theory was sound, the instrument was too complex for immediate field application.

5) Qualitative Feedback and Revision

The qualitative comments provided the roadmap for the necessary revisions to transition from Prototype 1 to the Final Model. Refining Digital Cognition: Validator 1 remarked, "*In the 'Digital Cognition' domain, the indicator for evaluating sources is too broad. Narrow it down to 'ability to distinguish opinion and fact in digital English texts'.*"

Response: The researcher revised the rubric. Instead of a general "Information Literacy" score, the final rubric specifically asks: "Can the student identify the author's bias and distinguish factual reporting from opinion in the selected online text?" This made the indicator measurable within an English reading comprehension task.

Addressing the Practicality Gap: Validator 2 provided crucial insight regarding the "Digital Divide," noting: "The language in the 'Socio-emotional' rubric is too academic. Teachers in remote areas might find it difficult to understand. Simplify the language to be more instructional (user-friendly). Add concrete examples."

Response: This feedback triggered a major overhaul of the linguistic presentation of the framework. Terms like "Synchronous Politeness Strategies" were simplified to "Politeness in Live Video Calls." Furthermore, a "Teacher's Guide" was added as an appendix to the framework, providing concrete scenarios. For example, for Domain 3 (Collaboration), the guide included a sample checklist: "Did the student use modal verbs (could, should) when giving suggestions to their friend?" This linked the abstract socio-emotional skill back to concrete linguistic markers (grammar), making it easier for EFL teachers to grade. The final result, post-revision, was a validated, user-friendly Digital Literacy Assessment Framework ready for empirical testing, boasting an average validity score that—after adjustment—would theoretically exceed the 3.5 threshold, though the primary achievement was the qualitative enhancement of its practicality.

3.2 Discussion

The findings of this study provide strong empirical support for the adoption of a structured approach to assessing digital literacy in the Indonesian EFL context, particularly when interpreted through the lenses of assessment literacy, multimodal pedagogy, and the structural constraints of Global South education systems. The results reveal a persistent misalignment between pedagogical innovation and assessment practices. Despite the increasing integration of digital tools in instruction, assessment remains predominantly paper-based (M = 4.65), confirming earlier observations that technology is often treated merely as a delivery mechanism rather than as a competence to be evaluated (Cahyono & Mutiaraningrum, 2016; Lie, 2020). This condition reflects what can be termed "assessment reductionism," wherein teachers focus narrowly on linguistic accuracy while neglecting broader communicative competencies in digital environments. Such practices result in construct under-representation, as defined by Messick (1996), because they fail to capture the full range of skills required for meaningful language use in contemporary contexts. Importantly, the findings suggest that this gap is less about limited access to technology and more about insufficient assessment literacy

among teachers, reinforcing the argument that 21st-century skills must be translated into operational criteria to become pedagogically meaningful (Griffin & Care, 2015) .

The tripartite framework developed in this study—comprising technical, cognitive, and socio-emotional domains—offers a theoretically grounded and contextually relevant response to these challenges. The inclusion of technical competence is supported by multimodality theory, which posits that meaning in digital texts is co-constructed through linguistic and semiotic resources (Kress, 2010). From this perspective, technical skills are not peripheral but foundational to effective communication, as limitations in navigating digital tools can obscure or distort linguistic output. Assessing technical competence, therefore, enhances construct validity by acknowledging the integral role of multimodal production in language learning. Moreover, it addresses concerns raised in prior research regarding the "invisible work" students perform in digital tasks, which often goes unrecognized in traditional assessment frameworks.

The cognitive domain extends the framework by incorporating critical digital literacy, particularly the ability to evaluate the credibility of online information. This is especially pertinent in contexts characterized by high internet usage but limited information literacy (Hidayat & Khalika, 2019). By embedding source evaluation into assessment criteria, the framework transforms reading from a passive activity into an active process of inquiry and verification. This aligns with Lankshear and Knobel's (2011) conceptualization of new literacies as involving not only new technical skills but also new epistemological orientations. Consequently, EFL assessment is reoriented toward fostering critical thinking, enabling students to engage with digital texts in a more reflective and informed manner.

The socio-emotional domain represents a particularly significant contribution, as it addresses the often-overlooked dimension of interpersonal communication in digital environments. The needs analysis indicated that teachers struggle to assess collaboration and online interaction, a challenge also noted in previous studies. By operationalizing socio-emotional constructs such as netiquette through observable linguistic features—such as politeness strategies and hedging—the framework bridges the gap between language proficiency and character education. This approach supports the view that digital platforms can facilitate, rather than hinder, the development of empathy and socially appropriate communication when guided by intentional pedagogical design (Ikhwan, 2019). In this way, communicative competence is expanded to include not only grammatical accuracy but also pragmatic and interpersonal effectiveness.

However, the study also underscores the importance of practicality in assessment design. The relatively low practicality score (M = 2.5) during initial validation highlights a common limitation in educational innovation: the development of theoretically robust but practically inaccessible tools. Feedback from experts indicated that overly technical language reduced the usability of the rubric, particularly for teachers with limited exposure to educational technology discourse. The subsequent revision, which simplified terminology and incorporated concrete examples, was essential in enhancing accessibility and addressing disparities associated with the digital divide (Madjid, 2002). These findings emphasize that validity in this context is inseparable from feasibility; an assessment framework must be both conceptually sound and practically implementable to be effective.

The implications of this study are significant for both policy and future research. For policymakers, the findings suggest that curricular reforms such as Kurikulum Merdeka should be accompanied by practical assessment tools that enable teachers to operationalize abstract competencies like critical thinking, collaboration, and digital literacy. Providing such tools can facilitate more consistent and meaningful implementation of curriculum goals. For future research, the validated framework offers a foundation for experimental studies examining its impact on student learning outcomes. Further investigation is also needed to explore how emerging technologies, including artificial intelligence, can support assessment processes—particularly in evaluating technical competencies—while preserving the role of teachers in assessing more nuanced socio-emotional dimensions.

4. CONCLUSION

This study addresses the "Assessment Lag" in Indonesian EFL instruction by constructing a multidimensional, theoretically grounded framework accompanied by a comprehensive set of validated indicators. These indicators provide a concrete roadmap for measuring "Digital Language Performance," shifting the focus from static grammatical knowledge to the functional use of English in complex digital environments. By integrating technical, cognitive, and socio-emotional domains, the framework ensures that assessment practices align with the authentic communicative demands of the 21st century.

Despite its contributions, this study has limitations that must be acknowledged. First, the practical validation phase was conducted within a specific geographic and institutional context, which may not fully capture the diverse range of technological infrastructure and digital literacy levels found across the entire Indonesian archipelago. Second, while the indicators were validated by experts and practitioners, the study did not include a long-term longitudinal analysis to measure the framework's impact on student proficiency over an extended period. Finally, the "practicality" of the framework—while addressed through iterative design—remains sensitive to the heavy administrative workloads currently faced by many Indonesian teachers. To build upon these findings, future research should focus on scaling the implementation of the framework across diverse socio-economic and regional contexts in Indonesia to ensure its broad applicability. Further studies could also investigate the development of digital assessment tools or AI-assisted platforms that automate the rubric-scoring process, thereby reducing the burden on teachers. Additionally, longitudinal research is recommended to evaluate how shifting from traditional to digital performance-based assessment influences students' long-term career readiness and global competitiveness. By continuing to refine these validated indicators, the Indonesian EFL community can ensure that assessment remains a bridge to the future rather than a relic of the past.

REFERENCES

Aydin, S. (2013). Teachers' perceptions about the use of computers in EFL teaching and learning: The case of Turkey. *Computer Assisted Language Learning*, 26(2), 214-233.

Cahya, D. (2020). *The digital divide in Indonesian education*. Jakarta Post.

Cahyono, B. Y., & Mutiaraningrum, I. (2016). Indonesian EFL teachers' familiarity with and opinion on the internet-based teaching of writing. *English Language Teaching*, 9(1), 199-208.

Dudeney, G., Hockly, N., & Pegrum, M. (2013). *Digital literacies*. Pearson Education.

Gee, J. P. (2017). *Literacy and education*. Routledge.

Gilster, P. (1997). *Digital literacy*. Wiley Computer Pub.

Godwin-Jones, R. (2018). Second language writing online: An update. *Language Learning & Technology*, 22(1), 1-15.

Griffin, P., & Care, E. (Eds.). (2015). *Assessment and teaching of 21st century skills: Methods and approach*. Springer.

Hafner, C. A. (2019). Digital literacies for English language learners. *TESOL Encyclopedia of English Language Teaching*.

Hidayat, D., & Khalika, R. (2019). Digital Divide in Indonesian Education: A Review of the Gap. *Journal of Education and Learning*, 11(2), 120-130.

Ikhwan, M. (2019). Socio-emotional learning in the digital age. *Indonesian Journal of English Education*, 6(1), 45-60.

Kemendikbud. (2020). *Panduan Pembelajaran Jarak Jauh*. Kementerian Pendidikan dan Kebudayaan.

Kirschner, P. A., & De Bruyckere, P. (2017). The myths of the digital native and the multitasker. *Teaching and Teacher Education*, 67, 135-142.

Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication*. Routledge.

Lamb, M. (2004). Integrative motivation in a globalizing world. *System*, 32(1), 3-19.

Lankshear, C., & Knobel, M. (2011). *New literacies: Everyday practices and social learning*. Open University Press.

Lie, A. (2020). Secondary education sector response to COVID-19. *ISEAS Yusof Ishak Institute*.

Madjid, N. (2002). *Indonesia: The challenge of change*. Brill.

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, 108(6), 1017-1054.

Ng, W. (2012). Can we teach digital natives digital literacy? *Computers & Education*, 59(3), 1065-1078.

Park, Y. (2012). A pedagogical framework for mobile learning. *International Review of Research in Open and Distributed Learning*, 12(2), 78-102.

Purnawarman, P., Susilawati, S., & Sundayana, W. (2016). The use of Edmodo in teaching writing in a blended learning setting. *Indonesian Journal of Applied Linguistics*, 5(2), 242-252.

Redecker, C. (2017). *European Framework for the Digital Competence of Educators: DigCompEdu*. Publications Office of the European Union.

Silviyanti, T. M. (2020). Looking into EFL teachers' perceptions on the use of technology in the classroom. *Journal of Language and Literature*, 10(2), 115-125.

Syahrul, S. (2022). Students'motivation In Learning English By Giving Warm Up Activities In Beginning Of Class. *Exposure: Jurnal Pendidikan Bahasa Inggris*, *11*(2), 214-236.

Son, J. B., Robb, T., & Charismiadji, I. (2011). Computer literacy and competency: A survey of Indonesian teachers of English as a foreign language. *CALL-EJ*, 12(1), 26-42.

Trilling, B., & Fadel, C. (2009). *21st century skills: Learning for life in our times*. Jossey-Bass.

Warschauer, M. (2011). *Learning in the cloud: How (and why) to transform schools with digital media*. Teachers College Press.

Zheng, B., Yim, S., & Warschauer, M. (2018). Social media in the writing classroom and beyond. *The TESOL Encyclopedia of English Language Teaching*.