

# A Narrative Review on the Effectiveness of Learning Evaluation to Sustain Professional Development among Indonesian Language Teachers

Marianus Supar Jelahut<sup>1</sup>

{suparjelahut@gmail.com<sup>1</sup>}

Universitas Katolik Indonesia Santu Paulus Ruteng, Jl.Jend.Ahmad Yani No.10 Kel.Tenda,  
Kec.Langke Rembong Kab, Manggarai, Nusa Tenggara Timur, Indonesia<sup>1</sup>

**Abstract.** This narrative review examines the effectiveness of learning evaluation in sustaining professional development among Indonesian language teachers. Using a narrative synthesis approach, the researcher critically analyzed fifty selected studies, theories, and policy documents addressing formative, authentic, and reflective evaluation practices. The review demonstrates that learning evaluation, when positioned as a process of continuous reflection rather than summative judgment, functions as a foundation for adaptive teacher growth and pedagogical renewal. It highlights the interrelated roles of formative assessment, authentic learning tasks, and reflective practice in shaping teachers' professional learning trajectories. The findings further indicate that effective evaluation practices depend not only on teacher competence but also on supportive institutional policies, digital readiness, and collaborative professional cultures. The review concludes that transforming evaluation into a catalyst for lifelong professional learning requires integrating evaluation theory, teacher agency, and policy coherence within a unified reflective system.

**Keywords:** learning evaluation, formative assessment, authentic assessment, reflective practice, teacher professional development

## 1 Introduction

Teacher professional development (TPD) has long been recognized as the backbone of educational reform and pedagogical renewal across the world. As global education systems shift toward learner-centered paradigms and competency-based curricula, the need to sustain teachers' professional growth becomes increasingly vital [1], [2], [3]. The quality of a nation's education system ultimately rests upon the learning and evaluative capacities of its teachers, as they translate policy and theory into meaningful classroom practice [4]. For Indonesia, where national curriculum changes and language education reforms are accelerating, developing sustainable frameworks of teacher professional development has become both a national priority and a scholarly necessity [5], [6], [7], [8].

Within this dynamic context, learning evaluation occupies a pivotal position as both a pedagogical and developmental process. Evaluation in learning is not merely an end-point

measurement; rather, it functions as an ongoing feedback mechanism that informs instruction, cultivates reflective teaching, and supports continuous improvement [9], [10]. Effective evaluation practices empower teachers to understand their own instructional decisions, identify gaps in student learning, and redesign pedagogical strategies accordingly [11], [12]. For language teachers in particular, learning evaluation represents a reflective mirror through which they can assess the alignment between linguistic competence, cultural understanding, and communicative performance [13], [14].

The Indonesian education system has undergone multiple reforms to strengthen teacher evaluation and performance standards. The “Merdeka Belajar” (Freedom to Learn) curriculum emphasizes teacher autonomy, formative assessment, and contextualized evaluation as essential to learning quality [6]. Yet, these innovations require teachers not only to implement assessment tools, but also to interpret and use evaluative evidence to refine their instructional practice [15], [16]. Consequently, teacher professionalism now depends on evaluation literacy as well as the capacity to design, interpret, and act upon evaluative data within authentic learning settings [17], [18].

While the importance of learning evaluation in supporting TPD has been widely discussed globally, there remains a lack of comprehensive conceptual synthesis addressing its role in Indonesian language education specifically. Studies on formative assessment, reflective teaching, and evaluation-based professional growth are dispersed across domains, leaving an unintegrated understanding of how evaluation functions as a sustainable driver for teacher learning [8], [19]. Furthermore, contextual constraints (such as limited teacher assessment literacy, unequal access to digital tools, and institutional barriers) continue to hinder evaluation implementation [20], [21]. The gap between evaluative policy frameworks and teachers’ real practices remains wide, creating a tension between compliance-based evaluation and developmental evaluation [7], [22].

Scholarly discourse has long recognized evaluation as both an art and a science of professional learning. The theoretical evolution of formative assessment since the works of Black and Wiliam (2009) has reframed evaluation from summative judgment to a collaborative learning process. Building on experiential learning theory [23] and reflective practice frameworks [24], evaluation is conceptualized as an iterative process of action, feedback, and adaptation. Such cyclical reflection forms the epistemological foundation of teacher growth; teachers learn by evaluating and re-evaluating their own instructional choices in the light of evidence and student outcomes [10], [25].

However, this theoretical transformation demands systemic support. Professional development initiatives that integrate evaluation as a learning mechanism, rather than as an external accountability tool, tend to generate more sustained change [1], [26]. Within Indonesian teacher training contexts, evidence shows that reflective evaluation workshops and peer-based assessment practices enhance teachers’ self-regulation, collaborative inquiry, and pedagogical adaptability [27], [28]. The interplay between evaluation and teacher reflection thus becomes a catalyst for professional autonomy and agency, essential for long-term professional growth.

At a deeper level, learning evaluation serves as the epistemic bridge between knowledge acquisition and professional transformation. The transformative learning theory [29] (Mezirow, 2015) positions critical reflection as the medium through which educators reinterpret their experiences and expand professional perspectives. In this sense, evaluation becomes more than

a procedural practice—it represents an intellectual habit of inquiry that nurtures reflective, evidence-informed, and adaptive teaching identities [30], [31] (Hobbs & Marks, 2020; Tractenberg, Chaterji, & Haramati, 2007). When teachers engage in structured reflection through evaluative cycles, they reconstruct their understanding of effective teaching, thereby deepening their professional competence [32], [33] (Vattøy & Gamlem, 2025; Hongyang, 2023).

In Indonesia's multilingual and socio-culturally diverse classrooms, this dimension of reflective evaluation is particularly crucial. Language learning is inherently contextual, shaped by learners' backgrounds, local culture, and communicative norms [34] [34] (Mitchell, Myles, & Marsden, 2019). Consequently, teachers' evaluation literacy must be sensitive to linguistic diversity and sociocultural variation. Studies highlight that when teachers design context-based evaluation instruments, the authenticity and relevance of learning increase significantly (Siti & Siti, 2019; Musaddat, Suryadi, & Suyanu, 2020) [35], [36]. Similarly, teacher-led innovations such as e-learning-based evaluations or whole-language learning assessments have demonstrated potential in fostering professional reflection and adaptive pedagogy (Andiopenta, 2023; Suhara, Firmansyah, & Permana, 2019) [37], [38].

Given these developments, the present narrative review seeks to provide a theoretical and conceptual synthesis of learning evaluation as a mechanism to sustain professional development among Indonesian language teachers. Specifically, it aims to: (1) identify and integrate dominant theoretical models of evaluation and professional learning; (2) interpret empirical evidence from Indonesian and global contexts; and (3) construct a conceptual framework linking evaluation practices to sustainable teacher growth. Through a narrative synthesis approach, this study interprets fifty carefully selected works spanning two decades of research, combining theoretical reflection and contextual critique. Thus, this article argues that learning evaluation, when understood as an ongoing, reflective, and participatory process, holds transformative potential for Indonesian language teachers' professional development. It calls for a paradigmatic shift; from evaluation as measurement to evaluation as learning, anchored in reflective inquiry, feedback literacy, and adaptive professional culture. In doing so, it contributes not only to theory building but also to reimagining evaluation as the sustaining pulse of professional development in the evolving landscape of Indonesian language education.

## **2 Literature Review**

Understanding the role of learning evaluation in sustaining teacher professional development requires revisiting its conceptual roots and the evolving pedagogical frameworks that shape it. Over time, evaluation has been reconceptualized from an administrative routine into a dynamic process of teacher learning. In this review study, the researcher explores how theoretical, empirical, and contextual strands converge to form a comprehensive perspective on evaluation as both a learning instrument and a professional development catalyst. Each subsection that follows aims to articulate key theoretical bases, relational constructs, and contextual nuances underpinning the interplay between learning evaluation and teacher growth.

## **2.1 Theoretical Foundations of Learning Evaluation**

The concept of learning evaluation has experienced significant evolution over the past decades. Initially viewed as a means to measure student achievement, evaluation has transformed into an integral component of the teaching and learning process (Stufflebeam & Coryn, 2014) [39]. The researcher recognizes that evaluation is not only about determining outcomes but also about generating insights for instructional improvement and reflective practice (Black & Wiliam, 2003) [40]. This shift toward formative and authentic assessment places evaluation at the core of pedagogical renewal, where the act of assessing becomes a shared learning experience for both teachers and students.

In this regard, formative evaluation functions as an ongoing cycle of evidence gathering and interpretation. It encourages teachers to make informed instructional decisions and refine their teaching strategies through real-time feedback (Gordon & McGhee, 2019) [12]. Within teacher education, this process mirrors what Kolb, Boyatzis, and Mainemelis (2014) describe as experiential learning, where knowledge is continuously reconstructed through experience, reflection, and adaptation. Learning evaluation thus becomes an experiential journey; teachers learn not only from students' progress but also from their own responses to learning data.

Building on reflective traditions, Schön (2013) characterizes the teaching act as a process of "reflection-in-action." The researcher adopts this lens to view learning evaluation as a self-dialogue: teachers interpret their practices while they are engaged in them, questioning their decisions and outcomes. When such reflection is structured and evidence-based, it leads to deeper professional insight. In this way, evaluation contributes to transformative learning, where teachers reshape their pedagogical identities through critical reflection (Mezirow, 2015) [29].

The researcher also identifies the importance of connecting evaluation with teacher change theories. Guskey (2002) explains that meaningful professional growth stems from teachers' experiences of improved learning outcomes, which in turn reshape their beliefs and practices. Similarly, Borko et al. (2000) emphasize that teacher education must treat evaluation as a situative process, embedded within classroom and community contexts. When evaluation is embedded in teachers' lived experiences, it nurtures professional agency; the ability to act reflectively, autonomously, and purposefully.

The theoretical synthesis presented in this review aligns with the notion that evaluation should not be an external judgment but a learning partnership between evidence and reflection. It draws together insights from formative assessment (Black & Wiliam, 2009) [40], experiential learning (Kolb et al., 2014), and reflective practice (Schön, 2013) to argue that teacher development flourishes when evaluation is integrated into the professional learning cycle. The researcher concludes that evaluation, when grounded in theory and reflection, becomes an indispensable foundation for sustainable teacher professional growth.

## **2.2 Learning Evaluation and Professional Development Nexus**

In this part of the review, the researcher focuses on how learning evaluation intertwines with teacher professional development (TPD). Across decades, studies have revealed that effective professional development requires teachers to engage in continuous learning cycles that include reflection, feedback, and classroom experimentation (Desimone, 2009; Darling-Hammond,

Hyler, & Gardner, 2017) [1], [2]. When these cycles are infused with evaluation, professional growth becomes not only reflective but also evidence-driven. Evaluation provides the interpretive lens through which teachers can analyze their progress and identify new pathways for improvement (Zeggelaar, Vermeulen, & Jochems, 2022) [4].

The essence of professional development lies in teacher transformation; a gradual shift in knowledge, attitudes, and beliefs (Kennedy, 2016) [3]. Through systematic evaluation, teachers can visualize how their teaching strategies impact learning outcomes and adjust accordingly (Goe, Biggers, & Croft, 2012) [26]. For instance, the integration of formative evaluation tools during workshops and coaching sessions has been found to reinforce teachers' feedback literacy and confidence (Ahadi, Bower, Lai, Singh, & Garrett, 2024; Fukubayashi & Fuji, 2025) [27], [41]. The researcher interprets such findings as evidence that evaluation itself can become a professional learning tool, especially when linked with mentoring and peer review systems.

Another dimension arises from the concept of self-assessment, where teachers and students alike become active participants in evaluating learning (Fitriyah, Wijayanti, & Purwaningrum, 2024). This participatory model of evaluation transforms professional development into a shared learning process. Teachers who engage their students in formative dialogue cultivate reflective habits that mirror their own professional inquiry (Shah & Mangrola, 2015) [42]. When feedback is reciprocal, classroom evaluation turns into collaborative professional learning rather than hierarchical judgment.

Moreover, evaluation serves as a bridge between teacher learning and institutional improvement. Research indicates that schools implementing continuous formative evaluation systems tend to foster more cohesive professional communities (Hobbs & Marks, 2020; Li & Gu, 2024) [28], [30]. These systems encourage teachers to share data, co-design rubrics, and co-reflect on learning patterns. The researcher argues that such collaboration reshapes evaluation from an individual task to a collective culture of inquiry. Through this process, professional learning communities evolve into reflective ecosystems that sustain long-term growth.

Professional development without evaluation risks becoming abstract or unmeasured. Evaluation without professional learning risks being punitive. Therefore, the nexus between the two represents a necessary equilibrium; a balance between accountability and learning. The researcher concludes that when evaluation is designed as a formative, reflective, and participatory practice, it enriches teachers' knowledge, agency, and commitment to continuous growth (Graham, 2023; Vattøy & Gamlem, 2025) [32], [43]. Through this synthesis, learning evaluation emerges as both the compass and the mirror of professional development, guiding teachers toward evidence-informed reflection and transformative practice.

### **2.3 Contextualization in Indonesian Language Education**

In the Indonesian educational landscape, learning evaluation represents both a professional challenge and a cultural adaptation process. The researcher acknowledges that Indonesia's diverse linguistic and socio-cultural realities demand context-sensitive approaches to assessment and teacher learning (Mappiasse & Sihes, 2014) [13]. Evaluation practices that might appear effective in global frameworks often require reinterpretation to align with the local curriculum, classroom diversity, and pedagogical traditions. Within this complexity, the

development of Indonesian language learning evaluation continues to evolve alongside broader curriculum reforms such as the Kurikulum Merdeka, which emphasizes authentic assessment, learner autonomy, and holistic learning outcomes (Amin & Rifa'i, 2024) [6].

Empirical studies have begun to illustrate how teachers adapt evaluation strategies to fit local learning goals. For instance, Andiopenta (2023) developed teaching materials based on **whole** language learning principles to improve evaluation literacy among Indonesian language teacher candidates. Similarly, Febriansyah and Syamsiyah (2024) proposed a contextual evaluation model that supports both formative and summative functions in language instruction. These innovations reflect how teachers are gradually moving from test-based evaluation toward reflective, student-centered approaches.

Yet, despite progress, evaluation reform faces persistent obstacles. Arrafii (2020) revealed that many teachers still equate evaluation with grading rather than learning improvement, resulting in mechanical implementation of assessments. Structural barriers, such as heavy administrative demands, limited training, and uneven access to professional support, continue to hinder transformative evaluation practices (Arsyad Arrafii, 2023; Sukenti & Tambak, 2020) [7], [20]. Furthermore, teachers' assessment literacy remains a crucial concern, particularly in distinguishing between assessment of learning and assessment for learning.

The researcher observes that a more sustainable path lies in integrating authentic and technology-enhanced evaluation. Studies have shown that digital-based tools, such as Socrative and e-learning platforms, facilitate more interactive and diagnostic feedback processes (Suhara, Firmansyah, & Permana, 2019) [38]. Authentic evaluation models that include project-based and text-based instruments also enhance students' engagement and teachers' reflective thinking (Musaddat, Suryadi, & Suyanu, 2020; Suwandi, Nurkamto, & Setiawan, 2019) [19], [36]. By connecting classroom realities with evaluative theory, these practices cultivate what the researcher calls reflective professional autonomy; the teacher's ability to make evaluative decisions grounded in both theory and lived classroom experience. Ultimately, contextualizing evaluation within Indonesian language education underscores the balance between global theory and local wisdom. This review study was introduced by the researcher to emphasize that professional development must grow organically from the classroom, where teachers themselves become the evaluators of their evolving pedagogical identities.

#### **2.4 Emerging Models and Innovations in Evaluation Practices**

The researcher further highlights that learning evaluation today is not confined to traditional frameworks; it is transforming through innovation, collaboration, and digital integration. Modern evaluation models emphasize the fusion of formative assessment literacy and technological empowerment, allowing teachers to adapt evaluation tools for more responsive and interactive learning contexts (Pastore, 2023; Suherman et al., 2025) [17], [18]. These innovations illustrate how evaluation is evolving from a static process into a dynamic, networked ecosystem of feedback and growth.

Several innovative models have gained prominence in recent years. One notable approach is lesson study, a collaborative professional learning strategy that integrates classroom observation, feedback, and reflective evaluation. Cheng and Lander (2024) explain that lesson study enables teachers to co-design and refine instructional strategies based on shared evaluative insights, thus reinforcing professional community and peer accountability. Similarly, Li and Gu (2024) demonstrate how formative assessment within continuing professional development (CPD) programs enhances teachers' self-regulated learning. Such models highlight that evaluation serves not only as measurement but as collaborative inquiry; a process through which teachers co-construct understanding and redefine their professional roles.

The researcher also notes an increasing trend toward digital evaluation innovations. The use of online formative tools and AI-supported feedback systems enables real-time data collection and analysis, promoting teacher adaptability and efficiency (Vattøy & Gamlem, 2025; Fukubayashi & Fuji, 2025) [32], [41]. These developments align with global movements toward digital professional development, where teachers use virtual platforms to exchange best practices and interpret feedback collaboratively (Ahadi et al., 2024). In Indonesia's emerging digital education landscape, these technologies are helping teachers overcome geographical and institutional barriers to professional growth.

Meanwhile, conceptual innovations also emerge from theoretical frameworks that link evaluation to teacher identity and motivation. Hongyang (2023) examined how professional development models in Asia incorporate reflective evaluation to sustain teacher self-efficacy. Atmowardoyo and Sakkir (2021) similarly suggested that theoretical models of language learning and evaluation should integrate contextual experience as a form of professional wisdom.

This review study was introduced by the researcher to emphasize that innovation in evaluation is not only technological but also cultural and philosophical. As teachers engage in continuous reflection through formative evaluation, they become researchers of their own classrooms; interpreting, adapting, and redefining what it means to teach effectively in changing contexts (Hobbs & Marks, 2020; Tractenberg, Chaterji, & Haramati, 2007; Werang et al., 2023) [22], [30], [31]. Ultimately, the researcher envisions a future where evaluation is seen as a living practice; a cycle of learning, unlearning, and relearning that keeps professional development alive and responsive.

### **3. Method**

This review study was introduced by the researcher through a narrative synthesis approach, an interpretive methodology that allows the integration of conceptual, empirical, and contextual evidence into a coherent theoretical understanding. The researcher selected this approach because it best accommodates the diversity of perspectives found within studies of learning evaluation and professional development. Unlike meta-analysis or statistical aggregation, the narrative synthesis emphasizes the interpretation and explanation of findings, enabling the researcher to identify conceptual relationships and construct theoretical linkages among the

reviewed works (Desimone, 2009; Stufflebeam & Coryn, 2014) [2], [39]. This qualitative logic aligns with the purpose of the present study—to build an interpretive model rather than to measure effect sizes.

The narrative synthesis was carefully adjusted to fit the reflective and conceptual nature of this topic. Learning evaluation and teacher professional development are complex phenomena shaped by theory, context, and practice. Therefore, the researcher adopted a method that values interpretive reasoning, allowing ideas to emerge inductively from the literature. Following the principles of narrative inquiry, this synthesis process involved identifying recurring themes, contrasting theoretical positions, and organizing them into evolving storylines that describe how evaluation contributes to sustainable teacher growth (Guskey, 2002; Kolb et al., 2014; Schön, 2013) [10], [23], [24]. This interpretive movement (from description to theorization) ensured that each selected source contributed meaningfully to the overall conceptual construction.

The references used in this study were not chosen randomly but derived from a deliberate and rigorous review process. The researcher explored hundreds of academic works (books, journal articles, and conference proceedings) sourced from the physical collections of a university library and online databases such as Taylor & Francis, Springer, Elsevier, and ERIC. From these, fifty references were systematically identified and approved as most relevant to the topic. The researcher employed a manual note-taking technique, recording essential theoretical arguments, definitions, and findings in a structured summary table. This process ensured that each citation included in this review directly supports the analytical framework and enriches the narrative coherence. Therefore, this approach was selected not only for its methodological suitability but also for its philosophical alignment with the researcher's interpretive stance. As a reflective inquiry into how evaluation sustains teacher learning, the narrative synthesis allows the research to evolve as a conceptual conversation; an academic dialogue among theories, practices, and lived experiences. Through this approach, the researcher upholds the scientific quality of this study while preserving the human dimension of scholarly reflection.

## **4. Findings and Discussion**

Based on the narrative synthesis conducted, the researcher identified several key thematic findings that collectively explain how learning evaluation contributes to sustaining professional development among Indonesian language teachers. The discussion in this section integrates conceptual, empirical, and contextual perspectives drawn from the reviewed literature to interpret the role of evaluation as both a reflective and transformative process. To ensure a coherent narrative flow, the researcher presents the findings in four interconnected themes that reflect theoretical, pedagogical, contextual, and innovative dimensions of learning evaluation in Indonesian language education. Each theme is analyzed interpretively within a reflective academic framework

### **4.1 Learning Evaluation as a Foundation for Teacher Professional Growth**

The researcher found that learning evaluation is far more than a technical procedure; it is a pedagogical foundation for teachers' continuous growth and reflective transformation. In many scholarly works, evaluation is described as the bridge between instructional practice and professional learning, positioning teachers not as passive implementers of standards but as active learners in their own classrooms (Black & Wiliam, 2003; Guskey, 2002; Schön, 2013)

[10], [24], [40].. The researcher interprets this as evidence that effective evaluation reshapes the teacher's role from evaluator to learner, where reflective action and evidence-based judgment promote sustained improvement. In this way, evaluation becomes a dialogic process between teaching experience, feedback, and adaptation (Kolb et al., 2014; Desimone, 2009; Kennedy, 2016) [2], [3], [23].

Globally, several studies demonstrate that evaluation systems which integrate formative and reflective principles can substantially influence teachers' professional agency. Black and Wiliam (2009) underline that formative assessment provides the necessary feedback loops to refine instructional decisions in real time. Darling-Hammond et al. (2017) similarly found that systematic evaluation tied to mentoring and peer reflection strengthens collective efficacy. Meanwhile, Ahadi et al. (2024) highlight that when teachers participate in evaluative workshops, their technological and pedagogical capacities co-evolve. The researcher regards these findings as showing that learning evaluation acts as an engine of continuous professional inquiry; a mechanism that allows teachers to monitor their growth while refining pedagogical quality.

However, the character and function of evaluation differ considerably across educational contexts. Based on Table 4.1 below, the researcher synthesized a comparative understanding of how international frameworks and Indonesian practices conceptualize the role of evaluation in teacher development. Internationally, evaluation tends to emphasize formative, collaborative, and feedback-driven mechanisms designed to foster reflection and autonomy (Black et al., 2006; Hobbs & Marks, 2020; Zeggelaar et al., 2022) [4], [9], [30]. In contrast, Indonesian perspectives, though increasingly reform-oriented, still exhibit a mixture of summative accountability and emerging reflective paradigms (Amin & Rifa'i, 2024; Arsyad Arrafii, 2023; Mulyati et al., 2023) [6], [8], [21].

**Table 1.** Comparative synthesis of international vs. Indonesian conceptualizations of evaluation for teacher growth

Focus / Concept	International Perspectives	Key Sources	Indonesian Perspectives	Key Sources
<b>Purpose of evaluation</b>	Viewed as formative, feedback-driven process to enhance instructional reflection and decision-making.	Black & Wiliam (2003, 2009); Guskey (2002); Desimone (2009)	Oriented toward accountability but shifting toward formative purpose in <i>Kurikulum Merdeka</i> .	Amin & Rifa'i (2024); Arrafii (2020); Fazalani (2025)
<b>Teacher's role</b>	Teachers as reflective practitioners who interpret evaluation evidence to improve pedagogy.	Schön (2013); Kolb et al. (2014); Kennedy (2016)	Teachers traditionally apply preset rubrics; reforms promote self-evaluation and inquiry.	Mulyati et al. (2023); Andiopenta (2023); Suwandi et al. (2019)
<b>Evaluation model</b>	Formative and authentic assessments emphasizing continuous feedback.	Black et al. (2006); Hobbs & Marks (2020); Morris et al. (2021)	Hybrid models combining summative scoring and authentic evaluation.	Amin & Rifa'i (2024); Mariyati & Putrayasa (2024); Suhara et al. (2019)
<b>Professional learning link</b>	Peer evaluation and feedback-based development sustain growth.	Ahadi et al. (2024); Borko et al. (2000); Darling-Hammond et al. (2017)	School-based reflection forums (MGMP) used for professional exchange.	Mulyati et al. (2023); Werang et al. (2023); Arsyad Arrafii (2023)
<b>Challenges / limitations</b>	Limited evaluative literacy; need for structural support.	Goe et al. (2012); Kennedy (2016); Zeggelaar et al. (2022)	Cultural emphasis on grades and administrative overload hinder reflection.	Arrafii (2020); Sukenti & Tambak (2020); Amin & Rifa'i (2024)
<b>Emerging trends</b>	Digital formative tools and coaching-based reflection.	Vatoy & Gamlem (2025); Fukubayashi & Fuji (2025); Cheng & Lander (2024)	Integration of e-learning evaluation tools and digital feedback.	Suhara et al. (2019); Mulyati et al. (2023); Andiopenta (2023)

Source: synthesized by the researcher from the 50 reviewed references

As reflected in Table 4.1, international frameworks present evaluation as a transformative learning mechanism, whereas Indonesian literature reveals an evolving but still transitional paradigm. The researcher notes that Indonesian teachers are progressively redefining their understanding of evaluation—not merely as grading, but as a reflective inquiry tool that informs

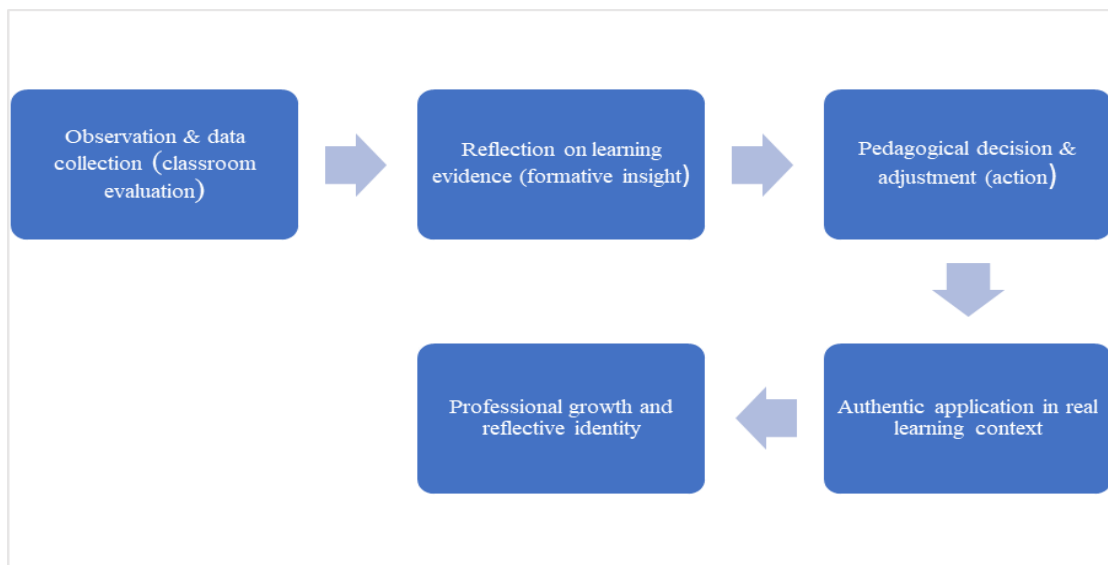
teaching strategies and self-assessment. Studies by Andiopenta (2023) and Mariyati and Putrayasa (2024) demonstrate how local innovations, such as evaluation based on whole-language learning and scoring techniques, gradually enhance teachers' professional reflection. Yet, as Arsyad Arrafii (2023) and Amin and Rifa'i (2024) emphasize, this shift requires stronger institutional and cultural support to develop evaluative literacy and sustain teacher motivation.

Synthesizing these conceptual and contextual strands, the researcher concludes that evaluation provides the intellectual and reflective foundation for professional growth. When teachers are empowered to evaluate their teaching as a learning process (supported by formative, authentic, and collaborative assessment) the boundary between teaching and learning dissolves. Teachers become learners of their own pedagogical experience, nurturing both competence and professional identity (Black et al., 2006; Guskey, 2002; Kennedy, 2016) [3], [10], [11]. This synthesis marks evaluation not merely as an administrative requirement but as a *transformative pedagogy* sustaining teachers' continuous professional development.

#### **4.2 Formative and Authentic Assessment as Catalysts of Reflective Practice**

The researcher observed that formative and authentic assessments are not only technical instruments but essential catalysts of reflective practice in teacher professional development. Across global and local studies, formative assessment is consistently portrayed as the process through which teachers learn about their own teaching by systematically gathering and interpreting feedback to adjust instruction [10], [40], [44]. Authentic assessment, in contrast, connects classroom evaluation with real-world tasks, encouraging teachers to reflect on how learning outcomes manifest in communicative and contextual settings (Suwandi et al., 2019; Mariyati & Putrayasa, 2024; Amin & Rifa'i, 2024) [6], [19], [45]. The researcher believes these assessment types are interdependent; formative assessment drives reflection through continuous feedback, while authentic assessment grounds reflection in meaningful pedagogical practice. Together, they serve as engines for teachers' professional inquiry and adaptive learning.

Building from these principles, Diagram 4.1 below conceptualizes the dynamic process through which formative and authentic assessment interact to stimulate teachers' reflective growth. The cyclical model—adapted and synthesized by the researcher; illustrates how evidence from classroom practice informs reflection, decision-making, and professional transformation.



**Fig. 1.** The Reflective Cycle of Formative and Authentic Assessment in Teacher Learning. Source: developed by the researcher based on synthesis from Black & Wiliam (2009), Schön (2013), and Suwandi et al. (2019).

As shown in Fig.1, the process begins when teachers engage in observation and evidence collection during lessons. This data, whether from student responses, peer feedback, or classroom performance, becomes the foundation for reflection. Through formative insight, teachers interpret these results to recognize strengths and learning gaps (Black et al., 2006; Li & Gu, 2024; Hidayat et al., 2023) [9], [11], [28], [46]. The next stage, pedagogical adjustment, involves transforming reflection into new instructional strategies. Authentic application then bridges theory and practice by situating evaluation in real communicative contexts where teachers can see the direct impact of their strategies (Mappiasse & Sihes, 2014; Mariyati & Putrayasa, 2024; Andiopenta, 2023) [13], [37], [45]. The final stage marks professional growth; teachers internalize reflection as an ongoing habit, forming what Schön (2013) called a “reflective identity.”

Globally, the literature converges on the notion that formative assessment nurtures a learning-oriented culture in schools. Black and Wiliam (2003) assert that when teachers use formative feedback consistently, they begin to adopt metacognitive habits similar to those they cultivate in students. Hobbs and Marks (2020) emphasize that extended formative evaluation promotes deeper engagement and professional collaboration. Similarly, Ahadi et al. (2024) report that teacher workshops centered on evaluative feedback enhance confidence and self-efficacy, particularly in digital settings. The researcher interprets these global insights as validation that formative evaluation is not only pedagogical but transformative; it changes the way teachers think about teaching itself.

In the Indonesian context, the integration of formative and authentic assessment has begun to reshape teachers’ perception of evaluation as reflective learning. Studies by Suwandi et al. (2019) and Amin and Rifa’i (2024) highlight that under the Kurikulum Merdeka reform, assessment is now viewed as an opportunity for teachers to examine their instructional methods

and adapt them to diverse learner needs. Similarly, Suhara et al. (2019) found that e-learning-based evaluation systems like Socrative enable teachers to provide immediate formative feedback, promoting reflective thinking through technology. However, challenges remain, such as limited digital literacy and inconsistent institutional support (Arsyad Arrafii, 2023; Mulyati et al., 2023; Werang et al., 2023) [8], [21]. The researcher thus argues that sustainable professional development depends on how well teachers internalize the reflective cycle illustrated in Diagram 4.1, integrating evaluation into their pedagogical consciousness.

Synthesizing across these studies, the researcher concludes that formative and authentic assessments function as twin pillars of reflective professionalism. Formative assessment provides structure to reflection, while authentic assessment ensures contextual relevance. Together, they transform evaluation from a static measurement activity into a dynamic learning process that enhances teacher self-awareness, agency, and competence (Black & Wiliam, 2009; Kennedy, 2016; Guskey, 2002) [3], [9], [10]. When teachers adopt this reflective cycle, they no longer perceive evaluation as an external obligation but as the internal heartbeat of professional growth.

### **4.3 Contextual and Policy Challenges in Indonesian Language Education**

The researcher identified that despite significant reform efforts, the implementation of effective learning evaluation within Indonesian language education remains constrained by contextual and policy-related challenges. One of the most pressing issues lies in the gap between national curriculum expectations and teachers' classroom realities. Although the Kurikulum Merdeka promotes formative and authentic assessment as pillars of modern pedagogy, many teachers still face difficulties in translating these principles into daily practice due to limited professional literacy, large class sizes, and restricted resources (Arsyad Arrafii, 2023; Amin & Rifa'i, 2024; Werang et al., 2023) [6], [8], [22]. Furthermore, the bureaucratic emphasis on summative assessment for accountability often overshadows formative feedback intended to guide teacher and learner growth (Fazalani, 2025; Mulyati et al., 2023; Ramly & Abduh, 2018) [14], [21], [47]. The researcher noted that this imbalance reflects a tension between compliance-based evaluation and developmental evaluation: a dilemma that continues to challenge teachers' professional autonomy.

Another crucial contextual constraint arises from the uneven distribution of technological and institutional support. As the education system moves toward digitalization, evaluation increasingly depends on teachers' digital competencies and infrastructure readiness. Studies by Suhara et al. (2019) and Ahadi et al. (2024) reveal that while some urban schools have adopted innovative e-learning platforms for assessment, rural and resource-limited regions remain constrained by inadequate facilities and weak internet connectivity. Consequently, professional development opportunities become unevenly accessed and unevenly effective (Vattøy & Gamlem, 2025; Hongyang, 2023) [32], [33]. The researcher observed that this technological divide also affects teachers' willingness to adopt reflective digital practices, as evaluation is often perceived as administratively burdensome rather than pedagogically empowering.

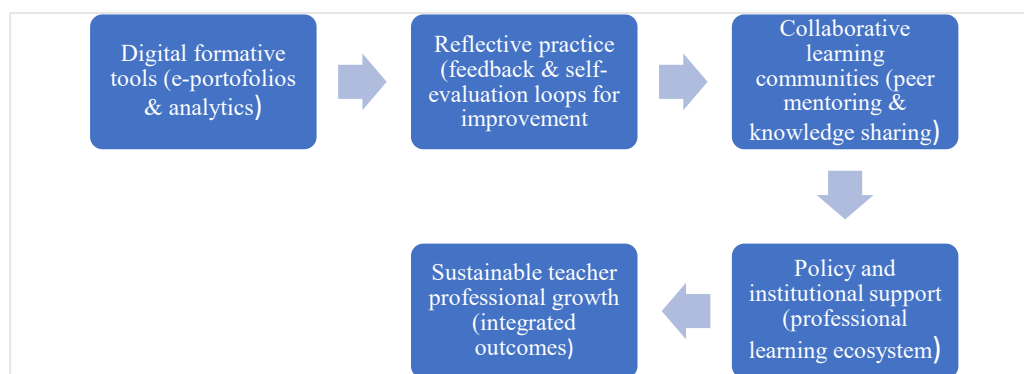
Policy coherence presents yet another layer of complexity. Despite national initiatives emphasizing teacher professionalization, the policies guiding evaluation are often fragmented across different ministerial and institutional levels (Sukenti & Tambak, 2020; Suwandi et al., 2019; Goe et al., 2012) [19], [20], [26]. This fragmentation leads to inconsistent implementation

standards and a lack of unified indicators for teacher evaluation literacy. The researcher argues that policy formulation must move beyond procedural standardization toward cultivating reflective and context-sensitive evaluation cultures. As Kennedy (2016) and Darling-Hammond et al. (2017) suggest, successful professional development emerges not from rigid frameworks but from systems that nurture teacher inquiry, experimentation, and reflection. Therefore, to sustain professional growth among Indonesian language teachers, evaluation must be re-envisioned as a collaborative process—one that aligns institutional policy, classroom context, and teacher agency in an integrated cycle of learning and reflection.

#### 4.4 Innovative Directions for Sustainable Professional Growth

The researcher views innovation in learning evaluation not merely as a technological upgrade but as a paradigm shift toward sustained professional growth. The emerging literature reveals that innovative directions must combine pedagogical, digital, and reflective dimensions in a synergistic framework. Globally, professional development has evolved toward models emphasizing inquiry-based evaluation, peer collaboration, and adaptive learning analytics (Ahadi et al., 2024; Kennedy, 2016; Zeggelaar et al., 2022) [3], [4], [27]. In the Indonesian context, innovation in evaluation needs to address both structural and cultural barriers by integrating digital tools, collaborative reflection, and contextual adaptation (Arsyad Arrafii, 2023; Mulyati et al., 2023; Suwandi et al., 2019) [7], [19], [21]. The researcher argues that sustainability in teacher growth arises when evaluation is no longer episodic but continuous; embedded within teachers' daily professional experience.

To visualize this transformative shift, Diagram 4.2 below presents the Innovative Evaluation Model for Sustainable Teacher Professional Growth, synthesized by the researcher from global and Indonesian perspectives. This model illustrates how innovation in learning evaluation can bridge the gap between teacher reflection, technological mediation, and institutional support to sustain ongoing professional learning.



**Fig.2.** Source: developed by the researcher based on synthesis from Kennedy (2016); Ahadi et al. (2024); Suwandi et al. (2019); Darling-Hammond et al. (2017); and Zeggelaar et al. (2022).

The first component of the model highlights the potential of digital formative tools such as e-portfolios, digital rubrics, and automated feedback systems to promote data-driven reflection.

These tools allow teachers to document learning progress and engage in self-assessment over time (Fukubayashi & Fuji, 2025; Suhara et al., 2019; Li & Gu, 2024) [28], [38], [41]. The researcher emphasizes that such digital innovation should not replace pedagogical judgment but complement it by visualizing learning evidence that can guide professional decisions. The second component, reflective practice, reinforces the formative cycle described earlier: teachers analyze collected data, identify patterns of success and challenge, and set personal learning goals (Schön, 2013; Black & Wiliam, 2009; Hobbs & Marks, 2020) [9], [24], [48].

The third element (collaborative learning communities) plays a crucial role in sustaining innovation. Research across contexts demonstrates that teachers' reflective practices thrive in collegial settings where feedback, mentorship, and shared inquiry occur regularly (Guskey, 2002; Borko et al., 2000; Prastikawati et al., 2024) [10], [16], [25]. In Indonesia, professional learning communities (Komunitas Belajar) have emerged as effective platforms for integrating evaluation and reflection, helping teachers align classroom practices with curriculum expectations while fostering mutual accountability (Mulyati et al., 2023; Werang et al., 2023; Arsyad Arrafii, 2023) [8], [21], [22]. The researcher believes these communities serve as social ecosystems that cultivate a culture of continuous evaluation-driven learning.

The fourth component concerns policy and institutional support, which function as enablers of sustainable innovation. Without consistent policies and administrative frameworks, innovative evaluation efforts risk remaining isolated classroom initiatives. Darling-Hammond et al. (2017) and Desimone (2009) both emphasize that sustainable professional development is achieved only when systemic structures (training programs, evaluation guidelines, and recognition mechanisms) align with reflective learning principles. Indonesian studies similarly indicate that coherent institutional support strengthens teacher motivation and evaluation literacy (Sukenti & Tambak, 2020; Suwandi et al., 2019).

Finally, the fifth component—sustainable teacher professional growth; represents the integrated outcome of the entire cycle. The researcher interprets sustainability not as permanence but as continuous adaptability: the ability of teachers to re-evaluate, redesign, and renew their professional practices in response to educational changes (Mezirow, 2015; Kennedy, 2016; Zeggelaar et al., 2022) [3], [29]. This ongoing renewal process enables teachers to maintain relevance and resilience within evolving pedagogical landscapes. In conclusion, the researcher asserts that innovative evaluation must be understood as a living, systemic process that merges technology, reflection, and collaboration. By adopting the model proposed in Diagram 4.2, Indonesian language teachers can transform evaluation into a dynamic driver of sustainable professional development. This integrated innovation, when supported institutionally and culturally, ensures that professional growth is not an external demand; but an internalized, self-sustaining practice of lifelong learning.

## **5. Conclusion**

This narrative review reaffirms that learning evaluation holds transformative potential in strengthening the professional development of Indonesian language teachers. Through conceptual synthesis and interpretive analysis, the researcher concludes that evaluation functions most effectively when it transcends its traditional measurement role and becomes an instrument of reflection, inquiry, and pedagogical self-renewal. When teachers engage in formative and authentic assessment processes, they do more than measure student outcomes; they cultivate reflective habits that sustain their own professional growth.

Nevertheless, this review acknowledges practical limitations within Indonesia's educational landscape. The prevailing emphasis on standardized testing, uneven access to digital resources, and inconsistent institutional support systems often prevent teachers from fully implementing reflective evaluation practices. These contextual challenges highlight the need for integrated policies and continuous professional development frameworks that empower teachers to embed reflective evaluation into daily instruction.

The study's scope was confined to a synthesis of theoretical and empirical literature, which limits the ability to generalize findings empirically. Therefore, future research should explore classroom-based applications of the proposed conceptual model through mixed-method or design-based approaches. Such studies could examine how reflective evaluation systems influence teaching efficacy, learner engagement, and professional identity formation over time. Ultimately, the vision derived from this review is to establish a reflective evaluation culture in which Indonesian language teachers learn from their own practice continuously. By uniting innovation, collaboration, and reflection, learning evaluation can evolve into a sustainable engine of professional transformation; empowering teachers not merely to teach, but to grow, adapt, and inspire.

## References

- [1] L. Darling-Hammond, M. E. Hyster, and M. Gardner, "Effective teacher professional development," *Learn. Policy Inst.*, 2017.
- [2] L. M. Desimone, "Improving impact studies of teachers' professional development: Toward better conceptualizations and measures," *Educ. Res.*, vol. 38, no. 3, pp. 181–199, 2009.
- [3] M. M. Kennedy, "How does professional development improve teaching?," *Rev. Educ. Res.*, vol. 86, no. 4, pp. 945–980, 2016.
- [4] A. Zeggelaar, M. Vermeulen, and W. Jochems, "Evaluating effective professional development," *Prof. Dev. Educ.*, vol. 48, no. 5, pp. 806–826, 2022.
- [5] S. A. Mahmud and M. I. M. Amin, "Teachers' Perspectives towards the Implementation of Assessment for Learning at University Level," *J. Univ. Raparin*, vol. 10, no. 1, pp. 239–258, 2023.
- [6] F. Amin and A. A. Rifa'i, "Profesionalisme Guru Dalam Asesmen Formatif dan Sumatif Pada Penerapan Kurikulum Merdeka," *Tadris J. Penelit. dan Pemikir. Pendidik. Islam*, vol. 18, no. 1, pp. 99–107, 2024.
- [7] M. Arsyad Arrafii, "Assessment reform in Indonesia: Contextual barriers and opportunities for implementation," *Asia Pacific J. Educ.*, vol. 43, no. 1, pp. 79–94, 2023.
- [8] M. A. Arrafii, "Towards formative assessment: exploring English teachers' conceptions and practices of assessment in Indonesia," 2020.
- [9] P. Black and D. Wiliam, "Developing the theory of formative assessment," *Educ. Assessment, Eval. Account. (formerly J. Pers. Eval. Educ.)*, vol. 21, no. 1, pp. 5–31, 2009.
- [10] T. R. Guskey, "Professional development and teacher change," *Teach. Teach.*, vol. 8, no. 3, pp. 381–391, 2002.
- [11] P. Black, R. McCormick, M. James, and D. Pedder, "Learning how to learn and assessment for learning: A theoretical inquiry," *Res. Pap. Educ.*, vol. 21, no. 02, pp. 119–132, 2006.
- [12] S. P. Gordon and M. W. McGhee, "The power of formative evaluation of teaching," in *Differentiated teacher evaluation and professional learning: Policies and practices for promoting career growth*, Cham: Springer International Publishing, 2019, pp. 15–35.
- [13] S. S. Mappiasse and A. J. B. Sihes, "Evaluation of English as a Foreign Language and Its Curriculum in Indonesia: A Review," *English Lang. Teach.*, vol. 7, no. 10, pp. 113–122, 2014.
- [14] R. Ramly and A. Abduh, "Exploring cognitive concepts in the national assessment of the Indonesian language," *New Educ. Rev.*, vol. 53, no. 3, pp. 142–152, 2018.
- [15] R. Hidayat, I. Sujadi, and B. Usodo, "Description of assessment: Assessment for learning and assessment as learning on teacher learning assessment," *J. Educ. Res. Eval.*, vol. 7, no. 4, pp. 653–661, 2023.
- [16] E. F. Prastikawati, M. A. Adeoye, and J. C. Ryan, "Fostering effective teaching practices: Integrating formative assessment and mentorship in Indonesian preservice teacher education," *Indones. J. Learn. Adv. Educ.*, vol. IJOLAE, pp. 230–253, 2024.

- [17] S. Pastore, "Teacher assessment literacy: A systematic review," in *Frontiers in Education*, vol. 8, Frontiers Media SA, 2023, p. 1217167.
- [18] A. Suherman, S. A. Surayya, A. Dwimaulani, and B. Setiawan, "A Study of Formative Assessment Literacy in Indonesia: Insight from Pre-Service EFL Teachers," *Scope J. English Lang. Teach.*, vol. 9, no. 2, pp. 856–866, 2025.
- [19] S. Suwandi, J. Nurkamto, and B. Setiawan, "Teacher Competence in Authentic and Integrative Assessment in Indonesian Language Learning," *Int. J. Instr.*, vol. 12, no. 1, pp. 701–716, 2019.
- [20] D. Sukenti and S. Tambak, "Developing Indonesian Language Learning Assessments: Strengthening the Personal Competence and Islamic Psychosocial of Teachers," *Int. J. Eval. Res. Educ.*, vol. 9, no. 4, pp. 1079–1087, 2020.
- [21] Y. Mulyati et al., "Pengoptimalan kemampuan evaluasi guru bahasa indonesia di era 5.0," *J. Pembelajaran Pemberdaya. Masy. (JP2M)*, vol. 4, no. 1, pp. 74–81, 2023.
- [22] B. R. Werang, I. M. Suarjana, K. K. Dewi, and S. I. Asaloei, "Indonesian language teachers' teaching performance and students' learning outcomes," *Int. J. Eval. Res. Educ.*, vol. 12, no. 3, pp. 1271–1277, 2023.
- [23] D. A. Kolb, R. E. Boyatzis, and C. Mainemelis, "Experiential learning theory: Previous research and new directions," in *Perspectives on thinking, learning, and cognitive styles*, Routledge, 2014, pp. 227–247.
- [24] D. Schön, "From technical rationality to reflection-in-action," in *Boundaries of adult learning*, Routledge, 2013, pp. 8–31.
- [25] H. Borko et al., "Teacher education does matter: A situative view of learning to teach secondary mathematics," *Educ. Psychol.*, vol. 35, no. 3, pp. 193–206, 2000.
- [26] L. Goe, K. Biggers, and A. Croft, *Linking Teacher Evaluation to Professional Development: Focusing on Improving Teaching and Learning. Research & Policy Brief. National Comprehensive Center for Teacher Quality*, 2012.
- [27] A. Ahadi, M. Bower, J. Lai, A. Singh, and M. Garrett, "Evaluation of teacher professional learning workshops on the use of technology-a systematic review," *Prof. Dev. Educ.*, vol. 50, no. 1, pp. 221–237, 2024.
- [28] J. Li and P. Y. Gu, "Formative assessment for self-regulated learning: Evidence from a teacher continuing professional development programme," *System*, vol. 125, p. 103414, 2024.
- [29] J. Mezirow, *Transformative learning. Challenging Educational Theories*, 2015.
- [30] L. Hobbs and G. Marks, "The power of formative evaluation: designing extended professional development," *Methodol. Approaches to STEM Educ. Res.*, vol. 1, p. 174.
- [31] R. E. Tractenberg, R. Chaterji, and A. Haramati, "Assessing and analyzing change in attitudes in the classroom," *Assess. Eval. High. Educ.*, vol. 32, no. 2, pp. 107–120, 2007.
- [32] K. D. Vattøy and S. M. Gamlem, "Navigating formative assessment as professional development in digital contexts: insights from teachers' experiences," *Teach. Dev.*, vol. 29, no. 2, pp. 343–360, 2025.

- [33] Y. Hongyang, "The evaluation mechanism of Russian Teacher Professional Development Model and its Inspiration for Teacher Education in China," *KANT*, vol. 2 (47), pp. 428–433, 2023.
- [34] R. Mitchell, F. Myles, and E. Marsden, *Second language learning theories*. Routledge, 2019.
- [35] P. J. I. Siti and I. Siti, "Context-based evaluation materials in elementary teacher education program: A developmental research," *J. Turkish Sci. Educ.*, vol. 16, no. 3, pp. 325–335, 2019.
- [36] S. Musaddat, I. Suryadi, and S. Suyanu, "The Validity of Text Based Evaluation Learning Instruments: Reconstruction of Character Evaluation Instrument Models in the Indonesian Language Learning," in *1st Annual Conference on Education and Social Sciences (ACCESS 2019)*, Atlantis Press, 2020, pp. 305–309.
- [37] A. Andiopenta, "The development of teaching materials of learning evaluation based on whole language learning for students at the Indonesian language and literature education study program," *Al-Ishlah J. Pendidik.*, vol. 15, no. 3, pp. 2935–2945, 2023.
- [38] A. M. Suhara, D. Firmansyah, and I. Permana, "Learning Evaluation Through E-learning Socratic in Language Learning," *J. Educ. Expert. (JEE)*, vol. 2, no. 2, pp. 115–122, 2019.
- [39] D. L. Stufflebeam and C. L. Coryn, *Evaluation theory, models, and applications*. John Wiley & Sons, 2014.
- [40] P. Black and D. Wiliam, "In praise of educational research': Formative assessment," *Br. Educ. Res. J.*, vol. 29, no. 5, pp. 623–637, 2003.
- [41] N. Fukubayashi and K. Fuji, "Development of a formative evaluation tool for coaching processes from the coachee and coach's perspectives," *J. Manag. Dev.*, 2025.
- [42] M. Shah and B. Mangrola, "Continuous professional development through feedback mechanism," *Voice Res.*, vol. 4, no. 1, 2015.
- [43] B. M. Graham, *A Theoretical Framework for Professional Development Program Evaluation and the Creation of an Associated Evaluation Tool*. North Carolina State University, 2023.
- [44] R. Morris, T. Perry, and L. Wardle, "Formative assessment and feedback for learning in higher education: A systematic review," *Rev. Educ.*, vol. 9, no. 3, p. 3292, 2021.
- [45] I. G. A. Mariyati and I. B. Putrayasa, "The Role of Scoring Techniques in Improving the Quality of Indonesian Language Learning Evaluation," *Bull. Sci. Educ.*, vol. 4, no. 3, pp. 387–395, 2024.
- [46] D. N. Hidayat, R. Diani, and S. Sutrisna, "The urgency of strengthening counseling guidance in remote education settings," *J. Psikol. Pendidik. dan Konseling*, vol. 8, no. 1, pp. 39–49, 2022.
- [47] R. Fazalani, "Peranan Penting Evaluasi Pembelajaran Bahasa Indonesia di Sekolah," *Innov. J. Soc. Sci. Res.*, vol. 5, no. 4, pp. 3937–3948, 2025.
- [48] R. Hobbs, *Digital and Media Literacy: A Plan of Action. A White Paper on the Digital and Media Literacy Recommendations of the Knight Commission on the Information Needs of Communities in a Democracy*. ERIC, 2010.