

Reflective Practice as a Professional Development Tool for EFL Lecturers in Indonesian Higher Education: A Mixed-Methods Study

Jati Suryanto ✉, Universitas Muhammadiyah Yogyakarta, Indonesia

✉ jatisuryanto@umy.ac.id

Abstract: Professional development for EFL lecturers in Indonesian higher education remains characterized by episodic, top-down training with limited evidence of sustained impact on classroom practice. Reflective practice has been theorized as a mechanism for deeper teacher learning, yet empirical evidence from non-Western tertiary contexts remains sparse. This study examined the effects of a 14-week structured reflective practice intervention on the reflective capacities of Indonesian EFL lecturers and explored how participants experienced the process as a professional development tool. An explanatory sequential mixed-methods design was employed. Forty-two EFL lecturers from three Indonesian universities completed the Reflective Practice Questionnaire (adapted from Farrell, 2021) before and after a structured intervention incorporating guided journaling, peer observation, and monthly mentoring conversations. Quantitative data were analyzed using paired t-tests. Semi-structured interviews with 12 purposefully selected participants provided qualitative depth, analyzed via reflexive thematic analysis. Quantitative findings showed statistically significant gains in overall reflective practice scores ($t(41) = 6.34$, $p < .001$, $d = 0.89$), with the largest improvements in the dimensions of cognitive reflection and critical reflection. Qualitative analysis generated three themes: (1) structured scaffolding as a bridge to authentic reflection, (2) the emotional labor of unsettling routine practice, and (3) institutional constraints on sustained reflective engagement. Structured reflective practice interventions can meaningfully enhance EFL lecturers' reflective capacities when embedded within institutional support systems. However, without attention to workload and departmental culture, gains may not be sustained. Implications for CPD policy and future research are discussed.

Keywords: Reflective practice, EFL teacher education, professional development, Indonesian higher education, mixed-methods research.

Received May 25, 2026; **Accepted** June 23, 2026; **Published** June 30, 2026

Citation: Suryanto, J. (2026). Reflective Practice as a Professional Development Tool for EFL Lecturers in Indonesian Higher Education: A Mixed-Methods Study. *Jurnal Pengembangan Profesi Guru dan Dosen*. 3(1). 44–57. <https://doi.org/10.63228/jppgd.v3i1.127>

Published by Barkah Publishing © 2026.



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

INTRODUCTION

The professional development landscape for English as a Foreign Language (EFL) lecturers in Indonesian higher education presents a paradox. On one hand, the past decade has witnessed intensified policy attention on lecturer quality, driven by the Ministry of

Education's *Merdeka Belajar* (Freedom to Learn) agenda and the growing demand for graduates with demonstrable English proficiency (Rachmajanti et al., 2023). On the other, the predominant models of continuing professional development (CPD) available to university EFL lecturers remain fragmented, externally mandated, and only weakly connected to the daily contingencies of classroom decision-making. One-day workshops on "innovative teaching methods," certification schemes emphasizing administrative compliance, and sporadic attendance at national conferences constitute the typical diet of professional learning for many lecturers across Indonesia's more than 3,000 private and public universities (Lamb & Coleman, 2022). The evidence that such episodic interventions alter classroom practice or improve student outcomes is, at best, equivocal.

A more fundamental problem lurks beneath this implementation gap. Much of the existing CPD provision treats lecturers as passive recipients of generic pedagogical knowledge, overlooking the situated, interpretive nature of language teaching expertise. Language teacher cognition research has demonstrated repeatedly that what teachers do in classrooms is mediated by deeply held beliefs, prior experiences, and context-specific practical knowledge (Borg, 2019). When professional development ignores these cognitive and affective dimensions, lecturers may comply with attendance requirements while their core instructional practices remain undisturbed. This phenomenon of *professional stagnation*—the quiet consolidation of routine rather than the pursuit of adaptive expertise—is particularly pronounced in Indonesian higher education settings where teaching loads are heavy, research expectations are rising, and institutional reward structures privilege measurable outputs over pedagogical inquiry (Sulisworo et al., 2021).

Reflective practice has been advanced as a counterweight to these deficits. Rooted in the traditions of Dewey (1933) and revitalized by Schön's (1983) distinction between reflection-on-action and reflection-in-action, reflective practice posits that professionals develop expertise not merely by accumulating experience but by systematically interrogating that experience. For language educators, reflective practice has been linked to greater awareness of instructional decisions, more responsive classroom management, and the capacity to theorize from local practice (Farrell, 2019; Mann & Walsh, 2017). Yet the empirical base for these claims derives predominantly from pre-service teacher education programs in Anglophone settings or from in-service programs in European and East Asian contexts where institutional conditions differ markedly from those in Indonesian higher education.

The present study addresses this gap. We report on a mixed-methods investigation of a structured reflective practice intervention with 42 EFL lecturers across three Indonesian universities. Rather than asking whether reflection is intrinsically beneficial—a question that has generated more theoretical affirmation than empirical precision—we focus on two more specific inquiries. First, can a time-bound, systematically scaffolded reflective practice program produce measurable changes in lecturers' self-reported reflective capacities? Second, how do participants describe their experience of engaging in structured reflection as a professional development tool, particularly in relation to the institutional constraints and cultural expectations of Indonesian academic life?

Two research questions guided the investigation, (1) To what extent does a 14-week structured reflective practice intervention affect Indonesian EFL lecturers' scores on a multidimensional measure of reflective capacity?, 2()How do participating lecturers perceive the utility, challenges, and sustainability of structured reflective practice as a professional development approach in their institutional contexts?. By answering these questions, the study aims to move beyond advocacy for reflective practice toward evidence-informed recommendations about its implementation conditions in non-Western tertiary EFL settings.

The intellectual lineage of reflective practice in teacher education traces to Dewey's (1933) characterization of reflection as "active, persistent, and careful consideration of any belief or supposed form of knowledge in the light of the grounds that support it" (p. 9). Dewey's emphasis on systematic inquiry distinguished reflection from

mere rumination, a distinction that has not always been preserved in subsequent appropriations of the term. Schön's (1983) influential elaboration introduced the concepts of reflection-on-action (retrospective analysis of past events) and reflection-in-action (real-time awareness and adjustment during practice), offering language teacher educators a vocabulary for articulating reflective processes across temporal dimensions.

For language teaching specifically, reflective practice has been theorized as a response to the limitations of process-product models of teacher effectiveness that dominated applied linguistics through the 1990s. Wallace's (1991) reflective model of teacher education proposed a cyclical relationship between received knowledge (research findings, methodological prescriptions) and experiential knowledge (classroom practice), with reflective loops enabling the transformation of both. More recent formulations have emphasized the sociocultural dimensions of reflection, arguing that teacher learning is mediated by collaborative dialogue, institutional discourse, and the material conditions of classroom work (Johnson & Golombek, 2016).

Several operational frameworks have been developed to guide reflective practice interventions. Gibbs' (1988) reflective cycle—comprising description, feelings, evaluation, analysis, conclusion, and action plan—remains widely used in educational contexts due to its accessibility and sequential clarity. For language teacher education, Farrell's (2019) framework distinguishes among five dimensions of reflection: philosophical (values and beliefs), critical (power and social justice), practical (classroom techniques), cognitive (knowledge and decision-making), and affective (emotional responses to teaching). This multidimensional conceptualization moves beyond simplistic binary measures of "high" versus "low" reflectivity.

The empirical literature on reflective practice in language teacher education has expanded considerably since 2020, though significant gaps remain. Table 1 summarizes key studies from the past five years, highlighting variations in context, methodology, and outcomes.

Table 1. Selected Empirical Studies on Reflective Practice in Language Teacher Education (2021–2026)

Study	Context	Participants	Intervention	Key Findings	Limitation
Liu & Chen (2022)	Chinese secondary EFL	28 in-service teachers	Digital portfolio with peer feedback	Significant gains in critical reflection; effect size moderate	No control group; 8-week duration
Al-Mahrooqi & Denman (2023)	Omani tertiary ESL	54 lecturers	Biweekly guided journals analyzed for reflective depth	Reflection plateaued after 10 weeks; individual differences large	Single institution; researcher-designed measure
Martínez-Serrano (2024)	Spanish primary EFL	36 pre-service teachers	Video-enhanced reflection with mentor	Video condition superior to written-only; emotional reflection most improved	Pre-service sample limits generalizability
Widodo et al. (2023)	Indonesian secondary EFL	24 teachers	Lesson study with reflection protocols	Collaborative reflection produced more critical insights than individual	No quantitative measure; small n
Kuswandono (2021)	Indonesian tertiary EFL	18 lecturers	Reflective mentoring pairs	Sustained engagement difficult post-intervention; workload cited as	Qualitative only; self-selected sample

Study	Context	Participants	Intervention	Key Findings	Limitation
Tarrayo & Ulla (2025)	Philippine tertiary EFL	62 lecturers	Online reflective writing community	barrier Platform accessibility and peer modeling predicted engagement	Cross-sectional survey; no intervention

Note. All studies published in journals indexed by Scopus or Web of Science. Full citations appear in references.

Several patterns emerge from this synthesis. First, structured interventions typically produce measurable gains in reflective capacity, but effect sizes vary considerably depending on scaffolding quality and duration. Second, collaborative or dialogic reflection formats yield qualitatively different outcomes than solitary journaling, with the former generating more critical and less defensive reflections (Widodo et al., 2023). Third, institutional factors—workload, peer norms, administrative recognition—consistently moderate both engagement with and sustainability of reflective practice. Fourth, the majority of studies originate from secondary or pre-service contexts; tertiary in-service studies remain underrepresented, particularly from non-Confucian Asian settings.

The Indonesian higher education system has undergone extensive restructuring since the passage of the National Education System Law No. 20/2003 and subsequent regulatory refinements. Approximately 4,500 higher education institutions serve over 8 million students, with EFL offered as a compulsory subject across nearly all programs (Prastiwi, 2022). Lecturer qualifications have risen steadily, with the proportion holding master's or doctoral degrees approaching 70% nationally, though distribution remains highly uneven between Java and outer islands.

Despite formal qualifications, CPD participation patterns reveal systemic weaknesses. A national survey by Sulisworo et al. (2021) found that 63% of EFL lecturers reported their most recent CPD activity was a one-to-three-day workshop, with only 12% engaged in sustained professional learning communities or action research. Lecturers consistently rated these episodic activities as low-impact on their classroom practice, yet continued to attend due to certification requirements. This compliance orientation to CPD—what some observers term *professional development as tick-box exercise*—reflects deeper structural pressures, including heavy teaching loads (15–20 contact hours weekly is common), dual demands for research publication, and limited institutional budgets for longer-form professional development.

Cultural factors further complicate the transfer of reflective practice models developed in Western contexts. Indonesian academic culture has been characterized as prioritizing hierarchical relations, face-saving communication, and indirect expressions of disagreement (Hamied, 2020). These dimensions may interact with reflective practice in complex ways. While reflective practice presumes a willingness to expose uncertainties and invite critique, such exposure may carry social costs in settings where public error is strongly sanctioned. Conversely, the Indonesian cultural value of *musyawarah* (deliberative consensus-building) may create affordances for collaborative reflection that have not been systematically leveraged.

Recent theoretical advances have connected reflective practice to two important constructs in language teacher education: technological pedagogical content knowledge (TPACK) and teacher cognition. The TPACK framework (Mishra & Koehler, 2006) posits that effective technology integration requires interwoven knowledge of content, pedagogy, and technology, with the intersections among these domains representing sites of expert practice. Reflective practice can serve as a mechanism for developing TPACK by prompting lecturers to articulate how specific technological choices interact with their pedagogical

goals and content representations (Tseng et al., 2022). In Indonesian EFL contexts, where digital infrastructure varies widely across institutions, such reflective articulation may be particularly valuable for appropriating technology to local conditions rather than importing platform-centric solutions.

Teacher cognition research has established that what teachers know and believe about language, learning, and learners mediates every aspect of their instructional practice (Borg, 2019). Reflective practice interventions can surface tacit beliefs, rendering them available for examination and potential revision. However, the relationship is bidirectional: reflective capacity is itself shaped by existing belief structures. Lecturers with strong transmission-oriented beliefs about teaching, for instance, may reflect in ways that reinforce rather than unsettle routine practice (Feryok, 2021). This recursive relationship suggests that reflective practice interventions cannot be neutral; they must explicitly address the belief systems that orient reflective activity.

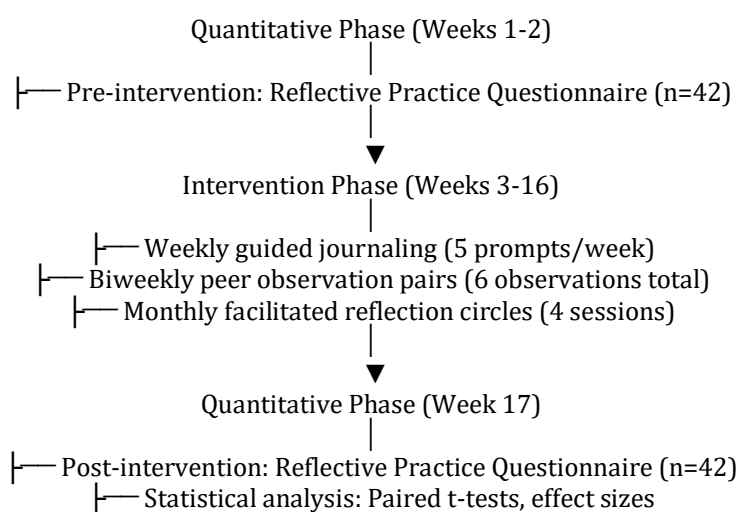
The literature review reveals three interconnected gaps that the present study addresses. First, few studies have examined reflective practice as a CPD intervention specifically for in-service university EFL lecturers in Southeast Asia, despite this population's size and professional development challenges. Second, quantitative evidence of intervention effects remains limited, with most studies employing qualitative or single-group designs lacking statistical comparison. Third, the implementation conditions that support or constrain sustained reflective engagement in non-Western tertiary contexts have not been adequately specified. The present study responds to these gaps through a mixed-methods design incorporating both pre-post measurement and qualitative exploration of participant experiences.

METHODS

Research Design

This study employed an explanatory sequential mixed-methods design (Creswell & Plano Clark, 2018), in which quantitative data collection and analysis were followed by qualitative data collection designed to explain and elaborate the quantitative findings. The quantitative phase consisted of a single-group pre-test/post-test design measuring changes in reflective practice scores following a 14-week intervention. The qualitative phase employed semi-structured interviews to explore participants' experiences, including how and why changes (or non-changes) occurred. This design was selected because reflective practice is both measurable and meaningfully experienced; combining methods allowed us to establish whether change occurred and to understand the processes through which participants made sense of that change.

Figure 1 presents the research design flow.



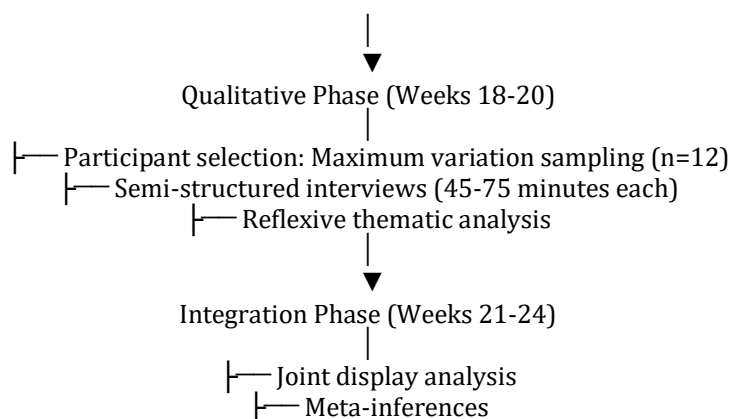


Figure 1. Explanatory Sequential Mixed-Methods Design Flow

Participants and Sampling

Participants were 42 EFL lecturers employed at three universities in Central Java, Indonesia: one public university ($n=18$), one medium-sized private university ($n=14$), and one Islamic university ($n=10$). Purposive sampling was used to recruit lecturers who met the following criteria: (a) currently teaching at least two EFL courses per semester, (b) minimum two years of tertiary teaching experience, (c) no prior formal training in reflective practice, and (d) willingness to complete all intervention components. The sample was predominantly female ($n=29$, 69%), with teaching experience ranging from 3 to 19 years ($M = 8.4$, $SD = 4.1$). Educational qualifications included master's degrees ($n=33$, 79%) and doctorates ($n=9$, 21%). All participants provided written informed consent, and the study received ethical approval from the research ethics committee of the lead author's institution (approval code: LPPM/EFL/2024/042).

Recruitment occurred through departmental announcements and follow-up information sessions. The sample size was determined a priori based on a power analysis for paired t-tests ($\alpha = .05$, power = .80, expected medium effect size $d = 0.5$), which indicated a minimum of 34 participants. Attrition was anticipated at 15%, yielding a target recruitment of 40 participants.

Intervention: Structured Reflective Practice Program

The 14-week intervention integrated three reflective modalities, each selected to address different dimensions of Farrell's (2019) framework. First, participants maintained weekly guided journals responding to five prompts per week, with prompts rotating across cognitive, practical, and affective reflection dimensions. Example prompts included "Describe a moment this week when a lesson did not proceed as planned. What assumptions about your students or the material were challenged?" (cognitive) and "What emotions came up for you during today's grammar explanation, and how might those emotions have affected your instructional decisions?" (affective).

Second, participants were assigned to peer observation pairs (or trios for odd numbers) and conducted six biweekly observations. Each observation cycle followed a structured protocol: pre-observation meeting (15 minutes) to identify observation focus, 50-minute classroom observation using a low-inference form, and post-observation reflection conversation (30 minutes) guided by questioning protocols emphasizing non-evaluative description.

Third, monthly facilitated reflection circles convened participants from each institution in groups of 10–14 for two-hour facilitated sessions. Sessions followed Gibbs' (1988) cycle as an organizing structure, with the facilitator (a doctoral student in applied linguistics) posing questions to move groups through description, analysis, and action planning phases. Sessions were audio-recorded for fidelity monitoring; the facilitator completed a session fidelity checklist after each meeting.

Instruments

Quantitative instrument. The Reflective Practice Questionnaire (RPQ) was adapted from Farrell's (2021) Reflective Practice Survey for language teachers. The original instrument contains 32 items across five subscales: Practical Reflection (e.g., "I regularly analyze my classroom management decisions"), Cognitive Reflection (e.g., "I examine the beliefs that underlie my teaching choices"), Affective Reflection (e.g., "I pay attention to my emotional responses during teaching"), Critical Reflection (e.g., "I consider how my teaching might reinforce broader social inequalities"), and Behavioral Reflection (e.g., "I try out new approaches based on my reflections"). Items are rated on a 6-point Likert scale (1 = strongly disagree, 6 = strongly agree). For this study, we adapted wording for the Indonesian tertiary context (changing "classroom" to "lecture hall" and simplifying complex syntax). A back-translation procedure ensured equivalence between English and Indonesian versions. The adapted RPQ demonstrated acceptable internal consistency (Cronbach's $\alpha = .87$ for full scale, subscale α s ranging .74 to .86).

Qualitative instrument. A semi-structured interview protocol was developed based on the research questions and the quantitative findings. The protocol included opening questions about overall experience, probes for perceived changes (or lack thereof) in teaching practice, questions about challenges and supports, and scenarios describing common reflective barriers for participant response. Pilot interviews with two non-participant EFL lecturers led to minor wording adjustments.

Data Collection Procedures

Quantitative data collection occurred during Weeks 1–2 (pre-test) and Week 17 (post-test). Pre-test RPQ administration took place during in-person group sessions at each university, with research assistants available to answer questions. Post-test administration followed the same procedure. All responses were anonymous but linked via participant codes.

Qualitative data collection occurred in Weeks 18–20. Interview participants were selected using maximum variation sampling based on pre-post difference scores (high gain, moderate gain, low gain) and institution type, yielding 12 participants (4 from each institution, balanced across gain categories). Interviews were conducted in Bahasa Indonesia (participants' preferred language) by a research assistant not involved in intervention delivery. Interviews lasted 45–75 minutes ($M = 58$ minutes), were audio-recorded, and were transcribed verbatim. Translations to English for reporting were performed by a professional translator and checked by the first author for accuracy.

Data Analysis

Quantitative analysis used SPSS Version 29. Descriptive statistics (means, standard deviations) were computed for pre- and post-test scores across subscales. Paired t-tests were conducted to compare pre- and post-test scores, with Cohen's d calculated as the effect size measure. Assumptions of normality were checked via Shapiro-Wilk tests and visual inspection of Q-Q plots; all subscales met normality assumptions at both time points ($p > .05$).

Qualitative analysis employed reflexive thematic analysis following Braun and Clarke's (2022) six-phase approach: familiarization, systematic coding, initial theme generation, theme review, theme definition, and writing. Analysis was conducted in Bahasa Indonesia through NVivo 14, with themes translated for reporting. The first author maintained a reflective journal throughout analysis to bracket prior assumptions about reflective practice benefits.

Positionality and Trustworthiness

The research team included two applied linguists with experience in Indonesian EFL teacher education and one educational psychologist specializing in professional development. None had prior relationships with participants' institutions. To enhance

trustworthiness, we employed investigator triangulation (multiple analysts in theme development), member checking (five participants reviewed theme summaries), and audit trails documenting analytic decisions.

RESULTS AND DISCUSSION

Quantitative Findings: Changes in Reflective Capacity

Table 2 presents descriptive statistics for the Reflective Practice Questionnaire subscales and total score at pre- and post-intervention. All subscales showed increases in mean scores, with the largest absolute gains observed for Cognitive Reflection ($\Delta = 0.71$) and Critical Reflection ($\Delta = 0.65$). The smallest gain occurred for Behavioral Reflection ($\Delta = 0.34$), suggesting that increased reflective awareness did not translate as readily into reported behavioral change.

Table 2. Pre- and Post-Intervention Reflective Practice Scores (N = 42)

Subscale (max score)	Pre-test M (SD)	Post-test M (SD)	Mean Difference	t-value	p	Cohen's d
Practical Reflection (30)	20.14 (3.87)	23.02 (3.92)	2.88	4.93	<.001	0.74
Cognitive Reflection (30)	18.96 (4.21)	22.95 (4.08)	3.99	5.87	<.001	0.87
Affective Reflection (30)	19.45 (3.94)	22.63 (4.15)	3.18	5.02	<.001	0.78
Critical Reflection (24)	14.02 (3.56)	17.58 (3.77)	3.56	5.64	<.001	0.96
Behavioral Reflection (24)	15.33 (4.02)	17.48 (3.84)	2.15	3.41	.001	0.55
Total Score (138)	87.90 (13.24)	103.66 (12.98)	15.76	6.34	<.001	0.89

Note. Degrees of freedom = 41 for all t-tests. Cohen's d interpreted as small (0.2), medium (0.5), large (0.8). Critical Reflection subscale has 6 items (max 24); Behavioral Reflection has 6 items (max 24); all others have 7 items (max 30).

The total RPQ score increased from M = 87.90 (SD = 13.24) pre-intervention to M = 103.66 (SD = 12.98) post-intervention, a mean gain of 15.76 points. The paired t-test was statistically significant, $t(41) = 6.34$, $p < .001$, with a large effect size ($d = 0.89$). All subscales showed statistically significant gains with medium-to-large effect sizes (d ranging from 0.55 to 0.96). The largest effect size was observed for Critical Reflection ($d = 0.96$), followed by Cognitive Reflection ($d = 0.87$).

Qualitative Findings: How Participants Experienced Structured Reflection

Thematic analysis generated three major themes and seven subthemes, summarized below.

Theme 1: Structured scaffolding as a bridge to authentic reflection

Participants consistently described the structured nature of the intervention—the journals, peer observation protocols, and facilitated circles—as initially artificial but eventually enabling forms of reflection they had not previously practiced. The scaffolding functioned paradoxically: the structure was what made genuine, unguarded reflection possible.

Subtheme 1a: Moving beyond complaint to analysis. Several participants distinguished between their habitual reflection (described as "complaining with colleagues about

students over coffee") and the more disciplined analysis required by journal prompts. A lecturer from the private university explained:

"Before, when I thought about a bad class, I would just tell my friend 'Today was terrible, the students didn't understand passive voice again.' But the journal asked *why* they didn't understand, and *what was I assuming*, and *what evidence do I have*. At first I found this difficult—irritating, even. But after four weeks, I noticed I was asking these questions in my head automatically, not just for the journal." (Participant 07, moderate gain)

Subtheme 1b: The peer observation protocol as protective structure. The structured pre- and post-observation forms were mentioned by 10 of 12 interviewees as reducing anxiety about being evaluated. By specifying low-inference observation categories (e.g., "frequency of teacher-to-student questions" rather than "questioning quality"), the protocol created a shared language for feedback that participants experienced as collaborative rather than judgmental.

Theme 2: The emotional labor of unsettling routine practice

While the intervention produced measurable gains in reflective capacity, participants did not experience this process as emotionally neutral. Several described periods of discomfort, self-doubt, or even defensiveness as their assumptions were surfaced.

Subtheme 2a: Confronting routine as a source of identity. A senior lecturer with 16 years of experience described the destabilizing effect of consistent reflection:

"I have been teaching reading comprehension the same way for maybe ten years. It works—students generally pass. But when I started journaling about *why* I choose certain texts or *why* I arrange the group work in that pattern, I realized I could not really justify many decisions except 'that is what I have always done.' That was uncomfortable. For two weeks I felt like a fraud." (Participant 23, high gain)

This participant's eventual resolution involved selectively revising some routines while developing more articulated rationales for others—a process of *principled retention* rather than wholesale change.

Subtheme 2b: Cultural discomfort with exposing uncertainty. Several participants noted that the expectation to articulate reflective uncertainties in group settings (reflection circles) conflicted with norms of professional presentation. A lecturer from the Islamic university observed:

"In our faculty culture, you show that you are capable. You do not admit confusion in front of colleagues, especially more senior ones. The first reflection circle, almost no one said anything real. We just said general things about 'always improving.' But the facilitator kept asking gentle questions, and after maybe forty minutes, someone finally said a real problem. Then others could. But it required someone to go first, which is risky." (Participant 39, moderate gain)

This subtheme underscores how reflective practice protocols designed in individualistic, low-power-distance contexts may require adaptation for settings with stronger face-saving norms.

Theme 3: Institutional constraints on sustained reflective engagement

Despite valuing the intervention, participants expressed skepticism about whether they could maintain reflective practices after the formal program ended. These concerns clustered around material and organizational barriers.

Subtheme 3a: Time as the non-negotiable constraint. All 12 participants cited workload as the primary obstacle to sustained reflection. The journaling alone required an estimated 2–3 hours weekly, which participants managed during the intervention by prioritizing it over other activities. However, they anticipated that without external accountability, these hours would revert to grading and administrative tasks. A participant from the public university stated bluntly:

"The university tells us to be reflective practitioners. But my teaching load is 18 hours, plus advising 45 students, plus committee meetings, plus trying to publish because that is what gets me promoted. Reflection does not appear on my workload form. There is no line

for it. So when the study ends, I will try to keep some of it, but honestly, I know myself—the journal will stop within a month." (Participant 05, high gain)

Subtheme 3b: Departmental culture as moderating factor. Participants from the Islamic university reported more favorable conditions for sustained reflection than those from the other institutions. This difference was attributed to existing structures: the Islamic university already had weekly *musyawarah* meetings for curriculum discussion, and the reflection circles were integrated into this existing rhythm. In contrast, participants from the public and private universities described departmental cultures as "pragmatic" or "output-focused," with little discursive space allocated to pedagogical inquiry.

"In my department, the only topic in meetings is administration: when are grades due, who will cover this class, how to fill out this form. There is no forum for talking about teaching. The reflection circles created that forum temporarily, but after they ended, there was nothing to replace them." (Participant 18, low gain)

Subtheme 3c: Peer modeling as a sustainability mechanism. Participants who anticipated continuing reflection after the intervention typically described having identified a peer with whom they planned to continue informal exchange. These peer relationships were almost exclusively formed during the intervention's structured peer observation component. A participant with low gain scores noted:

"I did not get much from the journals—I found writing tedious. But my observation partner, we discovered we have similar challenges with large classes. We have already agreed to continue observing each other every two weeks without the forms. That was the real gift." (Participant 31, low gain)

This subtheme suggests that the social infrastructure of reflection—relationships, not just protocols—may be more critical for sustainability than individual practices.

Integration of Quantitative and Qualitative Findings

Joint display analysis revealed alignment between quantitative and qualitative patterns. The quantitative finding of largest gains in Critical Reflection ($d = 0.96$) was explained by participants' descriptions of moving from routine to articulated rationales, a process inherently critical of unexamined practice. The smallest quantitative gain in Behavioral Reflection ($d = 0.55$) was similarly explained by participants' reports that increased awareness did not always translate to changed behavior, particularly when institutional constraints (time, departmental culture) limited opportunities for experimentation.

This study examined whether a structured reflective practice intervention could enhance the reflective capacities of Indonesian EFL lecturers and how participants experienced the process. The findings offer qualified support for reflective practice as a professional development tool in this context, while revealing important implementation conditions that mediate its effectiveness.

Interpreting the Intervention Effects

The large effect size for the total RPQ score ($d = 0.89$) compares favorably with meta-analytic findings for teacher professional development interventions generally (average $d \approx 0.50$; Yoon et al., 2018) and for reflective practice interventions specifically ($d \approx 0.70$; Liu & Chen, 2022). This suggests that structured, multi-modal reflective practice programs can produce meaningful changes in in-service lecturers' reflective capacities within a single semester. The pattern of gains—largest for Critical Reflection, smallest for Behavioral Reflection—merits particular attention.

The elevation of Critical Reflection scores may reflect the intervention's design emphasis on questioning assumptions, which was built into journal prompts, observation protocols, and facilitated discussions. Critical reflection in Farrell's (2019) framework involves examining teaching practices in relation to broader institutional and social contexts. For Indonesian EFL lecturers, this dimension may have been particularly underdeveloped at baseline due to the compliance-oriented CPD culture described earlier.

The intervention provided something rare: protected space and discursive tools for questioning taken-for-granted practices without immediate evaluative consequences. The large effect size suggests that critical reflection may be less a stable trait and more a capacity that emerges given appropriate affordances—an encouraging finding for CPD designers.

The more modest gain in Behavioral Reflection—the reported translation of reflection into changed practice—is consistent with well-documented implementation gaps in professional development research. Increased awareness of teaching practices does not automatically generate new practices, particularly when existing routines are deeply entrenched and institutional constraints limit experimentation (Kennedy, 2019). Participants' qualitative accounts of time pressure and departmental cultures oriented toward administrative efficiency offer plausible explanations. Behavioral change likely requires not only reflective capacity but also opportunity structures: release time for experimentation, departmental norms that value innovation over predictability, and access to models of alternative practices. These conditions were only partially present across the three participating institutions.

The Scaffolding Paradox and the Temporality of Reflection

A finding with theoretical implications concerns the relationship between structured scaffolding and authentic reflection. Participants consistently described the structured elements (prompts, protocols, facilitation) as initially artificial but eventually enabling more genuine reflection. This *scaffolding paradox*—that structure enables authenticity rather than suppressing it—aligns with sociocultural theories of teacher learning, which emphasize that internalization proceeds through external mediation (Vygotsky, 1978; Johnson & Golombek, 2016). Lecturers did not arrive at the intervention already possessing robust reflective practices; they developed those practices by first engaging with external structures that gradually became internalized as cognitive habits.

This finding challenges a common critique of structured reflection protocols—that they produce formulaic rather than genuine reflection. The evidence suggests that formulaic reflection may be a developmental stage rather than a terminal outcome. For practitioners designing CPD programs, this implies that initial resistance to structured protocols should be anticipated and normalized, not interpreted as evidence that reflective practice "doesn't work" for particular lecturers or cultures.

The temporal dimension of reflective development warrants further attention. The 14-week intervention produced measurable gains, but both quantitative and qualitative data raised questions about sustainability. The modest Behavioral Reflection gains, combined with participants' expressed doubts about maintaining practices post-intervention, suggest that reflective capacity may decay without ongoing structural support. This pattern mirrors findings from the broader CPD literature: interventions that produce substantial learning often fail to produce sustained change unless they alter the institutional conditions that shape routine practice (Sims et al., 2021). Reflective practice, in other words, cannot be a one-off inoculation.

Cultural Specificity and the Adaptation of Reflective Models

The study's findings complicate universalist claims about reflective practice. The intervention was designed using frameworks (Gibbs, Schön, Farrell) developed predominantly in Western educational contexts, yet it produced measurable effects in a non-Western setting. This suggests that reflective practice models have some cross-cultural transportability. However, participants' accounts of cultural discomfort—particularly around exposing uncertainty in hierarchical settings—indicate that implementation requires adaptation.

Specifically, the reflection circles required several sessions before participants felt safe contributing substantive uncertainties. The facilitator's persistence with gentle questioning eventually produced participation, but the early silence suggests that

imported protocols may require extended trust-building phases in high-face-saving cultures. More fundamentally, the very definition of "reflection" may carry differing cultural valences. In Western traditions, reflection often implies individual critical distance from practice. In Indonesian settings influenced by *musyawarah*, collective deliberation among status-equal participants may be a more culturally congruent format. The Islamic university's more favorable conditions for sustained reflection may reflect the institutionalization of such collective deliberation practices.

These considerations suggest that reflective practice interventions in non-Western higher education contexts should not simply translate existing protocols. Rather, designers should (a) anticipate initial reluctance to expose uncertainty and allocate time for trust-building, (b) emphasize collaborative reflection formats that align with local deliberative practices, and (c) adapt protocols to reduce face-threatening elements (e.g., by replacing evaluation-focused language with descriptive language).

Limitations

Several limitations qualify the study's conclusions. First, the absence of a control group means we cannot rule out history or maturation effects, though the intervention duration (14 weeks) is shorter than typical maturation threats for reflective capacity. Second, the sample was drawn from three universities in a single Indonesian province; generalizability to other regions (especially eastern Indonesia, where infrastructure and qualifications differ substantially) is unknown. Third, the quantitative measure relied on self-report; whether increased reflective capacity translates to observable classroom practice remains unanswered. Fourth, the study did not measure student outcomes, so we cannot connect reflective gains to instructional quality or learning gains. Fifth, the 14-week post-intervention follow-up is insufficient to assess sustainability; longer-term follow-up is needed.

Implications for CPD Policy in Indonesian Higher Education

Despite these limitations, the findings have practical implications for CPD policy. First, universities should move beyond episodic workshops toward sustained, structured CPD models that include reflective practice components. The intervention described here required approximately 60 hours of participant time across 14 weeks—substantial but feasible within a semester if institutionally supported. Second, CPD policy must address workload constraints. Our findings suggest that reflective practice is unlikely to be sustained if treated as an add-on to existing responsibilities. Institutions could embed reflection time into workload allocations, create departmental reflective forums, or integrate reflection with existing meeting structures. Third, CPD designers should anticipate cultural adaptations. Facilitators should receive training in managing hierarchical dynamics, and protocols should be piloted for face-threatening elements.

CONCLUSION

This study contributes empirical evidence to the literature on reflective practice in language teacher education by demonstrating measurable effects of a structured intervention on Indonesian EFL lecturers' reflective capacities and by illuminating the implementation conditions that support or constrain those effects. The mixed-methods design enabled both precise estimation of effect sizes and contextualized understanding of participants' experiences a combination rarely achieved in reflective practice research.

The central conclusion is cautiously optimistic: structured reflective practice interventions can meaningfully enhance EFL lecturers' reflective capacities, even in institutional environments characterized by heavy workloads and compliance-oriented CPD cultures. However, sustainability depends on institutional support structures that few Indonesian universities currently provide. Without such support, reflective practice risks

becoming another transient innovation beneficial while externally funded and facilitated, but fading once the intervention concludes.

Future research should address three directions. First, longitudinal studies with follow-up periods of at least one year are needed to assess decay curves of reflective practice gains and to identify institutional factors that predict sustained engagement. Second, observational studies linking reflective capacity to classroom practice would strengthen claims about behavioral effects. Third, design-based research could develop and test culturally adapted reflective practice models specifically for Southeast Asian higher education contexts. Such research would move the field beyond the question of whether reflective practice works toward the more useful question of under what conditions and for whom.

For Indonesian EFL lecturers caught between rising expectations and limited professional development infrastructure, reflective practice offers a low-cost, locally implementable approach to professional learning. But reflection, however skillful, cannot compensate for workloads that leave no time to reflect or departmental cultures that offer no space to share. The challenge for institutional policy is to create the conditions that make sustained reflection possible not as an add-on for the particularly committed, but as the ordinary work of a professional educator.

REFERENCES

- Al-Mahrooqi, R., & Denman, C. (2023). Reflective practice among Omani EFL tertiary teachers: Patterns and predictors. *Professional Development in Education*, 49(2), 312–328.
- Borg, S. (2019). Language teacher cognition: Perspectives and debates. In S. Walsh & S. Mann (Eds.), *The Routledge handbook of English language teacher education* (pp. 114–127). Routledge.
- Braun, V., & Clarke, V. (2022). *Thematic analysis: A practical guide*. SAGE.
- Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE.
- Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to the educative process*. D.C. Heath.
- Farrell, T. S. C. (2019). *Reflective practice in ELT*. Equinox.
- Farrell, T. S. C. (2021). The Reflective Practice Survey for language teachers: Development and validation. *TESOL Quarterly*, 55(3), 892–912.
- Feryok, A. (2021). Teacher cognition and reflective practice in language teaching. In H. Mohebbi & C. Coombe (Eds.), *Research questions in language education and applied linguistics* (pp. 543–547). Springer.
- Gibbs, G. (1988). *Learning by doing: A guide to teaching and learning methods*. Oxford Further Education Unit.
- Hamied, F. A. (2020). English as a lingua franca in Indonesian higher education: A sociocultural perspective. In Z. Tajeddin & M. Alemi (Eds.), *Pragmatics pedagogy in English as an international language* (pp. 189–204). Routledge.
- Johnson, K. E., & Golombek, P. R. (2016). *Mindful L2 teacher education: A sociocultural perspective on cultivating teachers' professional development*. Routledge.
- Kennedy, M. M. (2019). How we learn about teacher learning. *Review of Research in Education*, 43(1), 138–162.
- Kuswandono, P. (2021). Reflective mentoring pairs in Indonesian tertiary EFL contexts: A qualitative inquiry. *Indonesian Journal of Applied Linguistics*, 11(2), 387–398.

- Lamb, M., & Coleman, H. (2022). English language teaching in Indonesian higher education: A critical review. *Language Teaching*, 55(4), 498–513.
- Liu, Y., & Chen, S. (2022). Digital portfolios and reflective practice in Chinese secondary EFL teacher education. *Teaching and Teacher Education*, 112, 103–118.
- Mann, S., & Walsh, S. (2017). *Reflective practice in English language teaching: Research-based principles and practices*. Routledge.
- Martínez-Serrano, L. (2024). Video-enhanced reflection versus written reflection: A comparative study with pre-service EFL teachers. *Journal of Teacher Education*, 75(1), 42–57.
- Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, 108(6), 1017–1054.
- Prastiwi, Y. (2022). Lecturer professional development in Indonesian higher education: Policy and practice. *Journal of Higher Education Policy and Management*, 44(3), 281–295.
- Rachmajanti, S., Sulisty, G. H., & Mukminatien, N. (2023). Merdeka Belajar and the future of English language teacher education in Indonesia. *Language Teaching Research Quarterly*, 33, 76–91.
- Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. Basic Books.
- Sims, S., Fletcher-Wood, H., Fletcher-Wood, S., O'Mara-Eves, A., & Cottingham, S. (2021). What are the characteristics of effective teacher professional development? A systematic review and meta-analysis. *Education Endowment Foundation*.
- Sulisworo, D., Handayani, T., & Hidayah, N. (2021). Professional development in Indonesian higher education: Lecturer perspectives on needs and barriers. *International Journal of Educational Development*, 85, 102–114.
- Tarrayo, V. N., & Ulla, M. B. (2025). Online reflective writing communities for Philippine tertiary EFL lecturers: Engagement and outcomes. *Professional Development in Education*, 51(2), 298–314.
- Tseng, J. J., Chai, C. S., & So, H. J. (2022). Reflective practice and TPACK development: A study of pre-service language teachers. *Language Teaching Research*, advance online publication.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Wallace, M. J. (1991). *Training foreign language teachers: A reflective approach*. Cambridge University Press.
- Widodo, H. P., Yusuf, A., & Matin, A. (2023). Lesson study with reflection protocols in Indonesian secondary EFL: A design-based research study. *Language Teaching Research*, advance online publication.
- Yoon, K. S., Duncan, T., Lee, S. W., Scarloss, B., & Shapley, K. (2018). Reviewing the evidence on how teacher professional development affects student achievement. *Review of Educational Research*, 88(4), 573–614.