



PROCEEDINGS

International Conference on Education
Faculty of Tarbiyah | UIN Syekh Wasil Kediri
E-ISSN: 3025-6828

Volume 4, 2026

Pre-Service Indonesian Language Teachers' Needs and Readiness for a Local Wisdom-Based Pedagogical Chatbot

Zahra Alwi^{1*}; Khalidatun Nuzula²; Hani Atus Sholikhah³; Astrid Yulinda Putri⁴; Chrisnatama Tangguh Prasetyo⁵

^{1,2,3,4,5}Program Studi Pendidikan Bahasa dan Sastra Indonesia, Fakultas Keguruan dan Ilmu Pendidikan, Universitas Sriwijaya, Indonesia

Correspondence e-mail: zahra_alwi@fkip.unsri.ac.id

Abstract:

Pre-service Indonesian language teachers frequently struggle to translate Learning Outcomes (Capaian Pembelajaran) into concrete Learning Objectives and Trajectories (Tujuan and Alur Tujuan Pembelajaran), while South Sumatran local wisdom remains largely absent from their lesson designs and unsupervised generative-AI use further erodes critical digital literacy. This study reports a preliminary needs assessment and formative pilot evaluation of PAMONG-AI, a Socratic, local-wisdom-based pedagogical chatbot developed to accompany the Lesson Planning for Indonesian Language and Literature course at Universitas Sriwijaya. A sequential mixed-methods design combined a 20-item baseline competency test (N = 73) with a 51-item post-trial questionnaire (N = 80) covering technology acceptance, Socratic-interaction quality, local-wisdom integration, curriculum self-efficacy, usability, and dissemination readiness, supplemented by open-ended responses analyzed thematically. Results showed low baseline mastery (M = 33.8%, SD = 10.98), with the weakest areas being Deep Learning pedagogical principles (5%), learning-model syntax (9%), Modul Ajar structure (14%), and Kurikulum Merdeka/Kurikulum Berbasis Cinta distinctions (18%), precisely the gaps the chatbot was designed to close. After piloting the prototype, students reported favorable technology acceptance (M = 3.62/5), valued its Socratic questioning (M = 3.54/4) and local-wisdom integration (M = 3.54/4), and indicated improved curriculum self-efficacy (M = 3.53/4), while usability scored moderately (adapted SUS = 56.8/100), signalling concrete areas for refinement before expert validation. These findings validate the chatbot's design rationale and provide an empirical foundation for its next development iteration.

Keywords: Pedagogical Chatbot; Local Wisdom; Pre-Service Teachers; Technology Acceptance; Needs Assessment

Introduction

Pre-service Indonesian language teachers' pedagogical competence in designing lesson plans under Kurikulum Merdeka remains relatively low, while an unfavorable lecturer-to-student ratio severely limits the personalized feedback needed to close that gap, even after validated teaching materials have been made available (Alwi, Ernalida, & Lidyawati, 2020). The integration of South Sumatran local wisdom into lesson design is almost absent from student-produced teaching materials, even though prior work by this research group has repeatedly shown that local-culture-based materials strengthen literacy outcomes when deliberately scaffolded (Inderawati et al., 2022; Pratiwi, Purnomo, Alwi, & Zulkardi, 2023). At the same time, the unguided and widespread use of generative artificial intelligence among students weakens their critical digital literacy, as they increasingly copy AI output without verification (Alwi et al., 2026).

These conditions motivated the development of PAMONG-AI, a Socratic, local-wisdom-based pedagogical chatbot that accompanies the Lesson Planning for Indonesian Language and Literature course at the Indonesian Language and Literature Education Study Program, Universitas Sriwijaya. Rather than answering instantly, the chatbot is deliberately designed to question back, prompting students to articulate their school context before generating a lesson plan (Modul Ajar) aligned with Kurikulum Merdeka or the newer Kurikulum Berbasis Cinta, and to justify their choice of learning model, assessment type, and local-wisdom integration.

This study extends a line of needs-analysis research this group has previously conducted for other digital learning tools in South Sumatra, including an e-worksheet for poetry-writing skills (Edwin, Alwi, & Izzah, 2025) and a flipbook e-module integrating Palembang puppet heritage (Prihatmoko, Nurhayati, & Alwi, 2025), both of which found that need-based design substantially improves the perceived relevance and cultural authenticity of digital learning tools for Indonesian-language classrooms. Positioned within the conference theme of innovation and digitalisation for quality education, the present study contributes a rare combination: an objective baseline-competency measurement paired with a formative, multi-construct pilot evaluation of an actual AI prototype, rather than a purely intention-based survey. Its originality lies in triangulating what students objectively do not yet know with what they perceive after directly using the tool designed to teach it.

Before PAMONG-AI's features can be refined and submitted for expert validation, user needs must be systematically mapped so that its content and pedagogical approach genuinely address problems found in practice. This study therefore aims to describe (1) pre-service Indonesian language teachers' baseline competence in lesson-planning-related curriculum literacy, and (2) their perceptions and needs after piloting PAMONG-AI, covering technology acceptance, Socratic-interaction quality, local-wisdom integration, curriculum self-efficacy, usability, and dissemination readiness.

Methods

This study used a sequential descriptive mixed-methods design embedded within a larger research-and-development project following the Alessi and Trollip (2001) model. Two instruments were administered to fifth-semester students of the Indonesian Language and Literature Education Study Program, Faculty of Teacher Training and Education, Universitas Sriwijaya (Indralaya and Palembang classes), prior to the instruments' submission to expert validators. First, a 20-item multiple-choice baseline competency test covering curriculum literacy (Capaian Pembelajaran–Tujuan Pembelajaran–Alur Tujuan Pembelajaran), Bloom's taxonomy operational verbs, learning-model syntax, teaching competency frameworks, and assessment types was

administered through a gamified quiz platform (N = 73). Second, after piloting the PAMONG-AI prototype, students completed a 51-item structured online questionnaire (N = 80) comprising six closed Likert subscales — technology acceptance adapted from the Technology Acceptance Model (Davis, 1989), Socratic-interaction quality, local-wisdom integration, post-trial curriculum self-efficacy, an adapted System Usability Scale (Brooke, 1996; Bangor, Kortum, & Miller, 2008), and dissemination readiness adapted from Diffusion of Innovation theory (Rogers, 2003) — plus open-ended items on module-specific feedback and overall impressions.

Data were analyzed descriptively: item- and construct-level accuracy percentages for the competency test, and means and standard deviations for each Likert subscale. The adapted System Usability Scale was scored following Brooke's (1996) odd/even transformation and rescaled to a 0–100 range for an eight-item instrument. Open-ended responses were anonymized and analyzed thematically to identify recurring patterns in perceived benefits, concerns, and improvement requests.

Results and Discussion

Findings

Baseline Pedagogical Competence

Across the 73 participants who completed the baseline competency test, overall class accuracy was 34%, with an individual mean of 33.8% (SD = 10.98, median = 35%); no participant scored above 55%. Table 1 summarizes accuracy by construct cluster.

Table 1. Baseline Competency Test Accuracy by Construct Cluster (N = 73)

Construct Cluster	Item(s)	Mean Accuracy
General purpose of lesson planning	1	82%
Applied learning-objective (TP) formulation	19	60%
Model-context application	14, 20	52%
Bloom's taxonomy operational verbs (KKO)	4, 5, 6, 7	48%
Teacher competency legal basis	17	41%
Assessment type identification	18	29%
CP-TP-ATP structural literacy	2, 3	32%
Basic Teaching Skills (KDM) linkage	16	22%
Kurikulum Merdeka vs. Kurikulum Berbasis Cinta	8, 9	18%
Modul Ajar document structure	10	14%
Learning-model syntax recognition	12, 13, 15	9%
Deep Learning pedagogical principles	11	5%

Source: Baseline competency test results, PAMONG-AI pilot (2026).

The pattern shows that students could reliably reason about the general purpose of planning and apply a plausible objective or model in a described scenario, but consistently failed on recently regulated or highly structural knowledge: the three Deep Learning principles introduced by Permendikdasmen No. 13/2025 were recognized by only 5% of participants, the specific step sequence of named learning models by 9%, the required components of a Modul Ajar by 14%, and the jurisdictional distinction between Kurikulum Merdeka and Kurikulum Berbasis Cinta by 18%. This mirrors the competence-gap pattern this research group has previously documented when analyzing needs for other digital Indonesian-language tools (Edwin, Alwi, & Izzah, 2025; Prihatmoko, Nurhayati, & Alwi, 2025): pre-service teachers tend to hold serviceable general pedagogical intuition while lacking the precise, regulation-anchored literacy that a written lesson plan actually requires.

Technology Acceptance of PAMONG-AI

Table 2 presents descriptive statistics for all closed-ended subscales collected after students piloted the prototype (N ranged 71–80 per item due to unanswered items).

Table 2. Post-Trial Questionnaire Subscale Descriptive Statistics

Subscale	Scale	Mean	SD
Perceived Usefulness (TAM)	1–5	3.69	0.49
Perceived Ease of Use (TAM)	1–5	3.53	0.54
Attitude Toward Use (TAM)	1–5	3.63	0.49
Behavioral Intention (TAM)	1–5	3.67	0.49
Socratic Interaction Quality	1–4	3.54	0.51
Local Wisdom Integration	1–4	3.54	0.51
Curriculum Self-Efficacy (post-trial)	1–4	3.53	0.51
Usability (adapted SUS)	0–100	56.8	—
Dissemination Readiness	1–4	3.61	0.50

Source: PAMONG-AI post-trial questionnaire (2026).

All four Technology Acceptance Model subscales scored above the scale midpoint, with Behavioral Intention ($M = 3.67$) and Perceived Usefulness ($M = 3.69$) rated highest, consistent with the TAM proposition that perceived usefulness is the strongest proximal driver of continued-use intention (Davis, 1989). Perceived Ease of Use scored comparatively lower ($M = 3.53$), foreshadowing the usability concerns detailed below.

Socratic Interaction and Local Wisdom Integration

Students rated the five-stage Socratic consultation flow favorably ($M = 3.54/4$), agreeing that it prompted reflection before providing guidance and that its questioning pattern was distinguishable from a generic AI chatbot that answers immediately. Local-wisdom integration was rated similarly ($M = 3.54/4$). Notably, in the open-ended item asking which additional local wisdom should be added, 71 of 80 respondents proposed specific items — including Rejung Semende, Tari Gending Sriwijaya, Bekarang Iwak, and Senjang — indicating that students engaged critically and generatively with the cultural-integration feature rather than treating it as decorative, and directly extending the cultural corpus this group has documented in earlier work on Besemah folklore (Pratiwi, Purnomo, Alwi, & Zulkardi, 2023).

Curriculum Self-Efficacy, Usability, and Dissemination Readiness

Post-trial curriculum self-efficacy averaged 3.53/4 across CP-TP-ATP differentiation, KKO-based objective formulation, model selection, Basic Teaching Skills linkage, and assessment differentiation. This self-reported confidence is considerably higher than the objective baseline accuracy of 33.8% observed before the pilot, a discrepancy that should not be read as a validated learning gain — no matched post-test was administered in this cycle — but rather flags a self-assessment gap worth verifying with an objective post-test in the next research cycle.

The adapted eight-item System Usability Scale yielded a score of 56.8 out of 100. Using Bangor, Kortum, and Miller's (2008) adjective-rating benchmarks, where scores near 68 are considered average, this result sits in the "OK" range but below average, converging with open-ended feedback requesting simpler navigation, a more complete glossary, and clearer access instructions. Dissemination readiness was rated positively overall ($M = 3.61/4$); students agreed that practicing teachers, not only pre-service ones, would benefit from the tool ($M = 3.72/4$) and that it could still function in schools with limited facilities ($M = 3.53/4$), while simultaneously agreeing that further development is needed before wider dissemination ($M = 3.62/4$) — an internally consistent signal of cautious, conditional readiness rather than uncritical enthusiasm.

Qualitative Themes

Thematic analysis of the open-ended responses ($n = 44\text{--}74$ non-empty responses per item) surfaced three recurring themes. First, perceived efficiency and usefulness: respondents described the tool as “mempercepat pekerjaan” (speeding up the work) and contextually relevant to their region. Second, critical AI-literacy concern: several respondents explicitly worried about over-reliance, with one noting the risk that “guru menjadi terlalu bergantung pada AI sehingga kurang melakukan analisis dan penyesuaian terhadap karakteristik peserta didik” (teachers becoming too dependent on AI and under-analyzing learner characteristics), echoing this group’s prior caution about unguided generative-AI use. Third, concrete refinement requests: students asked for a more complete glossary, more varied examples, and assurance that AI-generated output remains editable rather than final — directly actionable input for the next development iteration before expert validation.

Product Access and Interface

PAMONG-AI is deployed as a web application accessible through the principal investigator’s academic portfolio site, providing a stable, citable point of access for participants and future users (Figure 1). The prototype’s landing page orients first-time users to its six modules — Ruang Konsultasi, Konsultan CP/ATP, Studio Modul Ajar, Perpustakaan Kearifan Lokal, Cermin Literasi Digital, and Glosarium — before they proceed to the Socratic consultation stage (Figure 2).

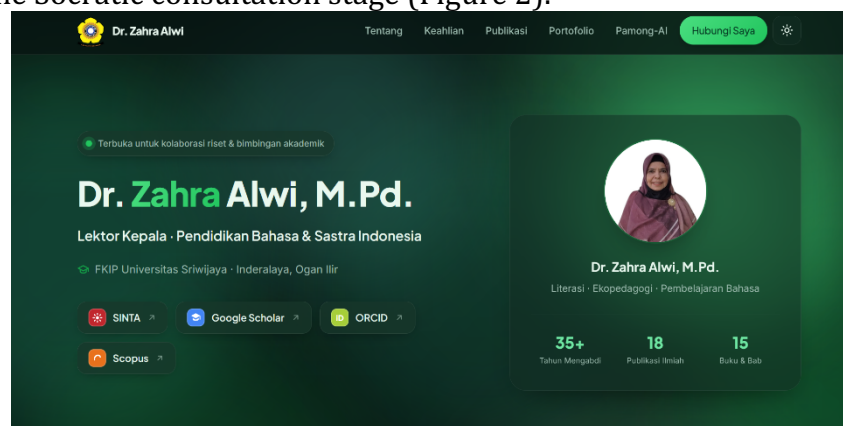


Figure 1. Access point to the PAMONG-AI prototype via the principal investigator’s academic portfolio site.

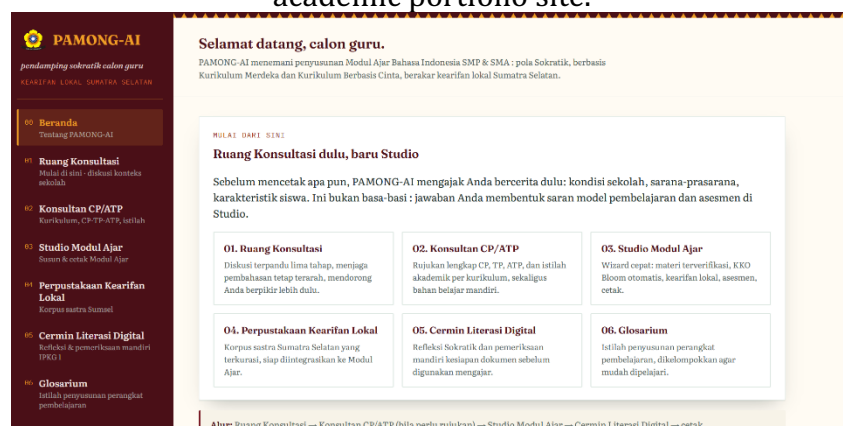


Figure 2. PAMONG-AI landing page introducing its six pedagogical modules.

Analysis

Triangulating the two data sources clarifies the chatbot's design rationale. The baseline test isolated exactly the knowledge areas — Deep Learning principles, learning-model syntax, Modul Ajar structure, and Kumer/KBC distinctions — that PAMONG-AI's Konsultan CP/ATP and Studio Modul Ajar modules are built to scaffold, while the post-trial questionnaire shows students recognized and valued exactly these features (technology acceptance and self-efficacy subscales both above the scale midpoint). This alignment supports the R&D principle that instructional technology should be designed from an empirically mapped deficit rather than an assumed one (Alessi & Trollip, 2001), and is consistent with this research group's established practice of grounding digital tool design in explicit needs analysis (Edwin, Alwi, & Izzah, 2025; Prihatmoko, Nurhayati, & Alwi, 2025).

At the same time, the usability score and open-ended feedback jointly indicate that acceptance is not yet unconditional: ease-of-use trailed usefulness in the TAM subscales, the adapted SUS score fell below the conventional average benchmark, and qualitative responses converged on interface simplification and glossary completeness as concrete priorities. Read together with the self-efficacy/accuracy discrepancy noted above, these findings argue for two next steps ahead of expert validation: (1) an interface and glossary refinement cycle directly responsive to the qualitative feedback collected here, and (2) inclusion of a matched post-test in the next pilot cycle so that self-reported confidence gains can be cross-validated against objective competency change, consistent with recommended practice in AI-literacy research (Alwi et al., 2026).

Conclusion

This study mapped a substantial baseline competence gap among pre-service Indonesian language teachers ($M = 33.8\%$ accuracy), concentrated in the newest and most structurally demanding areas of curriculum literacy — Deep Learning principles, learning-model syntax, Modul Ajar structure, and the Kurikulum Merdeka/Kurikulum Berbasis Cinta distinction. After piloting PAMONG-AI, a Socratic, local-wisdom-based pedagogical chatbot designed to address precisely these gaps, students reported favorable technology acceptance, valued its reflective questioning style and cultural integration, and reported improved curriculum self-efficacy, while usability and interface feedback identified concrete, actionable priorities for the next development cycle. These results validate the chatbot's core design rationale and provide an empirical, user-grounded foundation before the prototype proceeds to expert validation by material, media, pedagogical, and cultural validators. Future research should incorporate a matched objective post-test to verify whether the self-reported confidence gains observed here correspond to measurable competency change, and should track how the module-specific refinements requested by students are subsequently addressed in the revised prototype.

References

- Alessi, S. M., & Trollip, S. R. (2001). *Multimedia for learning: Methods and development* (3rd ed.). Allyn & Bacon.
- Alwi, Z., Ernalida, & Lidyawati, Y. (2020). Kepraktisan bahan ajar perencanaan pembelajaran berbasis pendidikan karakter dan saintifik. *Fon: Jurnal Pendidikan Bahasa dan Sastra Indonesia*, 16(1), 10–21. <https://doi.org/10.25134/fjpbsi.v16i1.2312>
- Alwi, Z., Ernalida, Nuzula, K., Gunaika, R., Salsabila, N., & Aulia, P. P. (2026). Pelatihan guru dalam peningkatan literasi siswa melalui pemanfaatan e-modul berbasis

- ekopedagogi berbantuan kecerdasan buatan. *Jurnal Pengabdian Masyarakat Ilmu Pendidikan*, 5(1), 106–116. <https://doi.org/10.23960/jpm-ip.vol.5i.1.1565>
- Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An empirical evaluation of the System Usability Scale. *International Journal of Human–Computer Interaction*, 24(6), 574–594. <https://doi.org/10.1080/10447310802205776>
- Brooke, J. (1996). SUS: A “quick and dirty” usability scale. In P. W. Jordan, B. Thomas, B. A. Weerdmeester, & I. L. McClelland (Eds.), *Usability evaluation in industry* (pp. 189–194). Taylor & Francis.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340. <https://doi.org/10.2307/249008>
- Edwin, M., Alwi, Z., & Izzah. (2025). Developing interactive E-LKPD with Wizer.Me: A needs analysis for enhancing poetry writing skills in junior high school students in Palembang. *Al-Ishlah: Jurnal Pendidikan*, 17(2), 1989–2005. <https://doi.org/10.35445/alishlah.v17i2.6811>
- Inderawati, R., Amrullah, Mirizon, S., Purnomo, M. E., Ernalida, Indrawati, S., Shilviany, S., & Alwi, Z. (2022). South Sumatra local culture-based teaching materials training for vocational school language teachers. *English Franca: Academic Journal of English Language and Education*, 6(1), 65–82. <https://doi.org/10.29240/ef.v6i1.3665>
- Permendikbudristek No. 12 Tahun 2024 tentang Perubahan atas Peraturan Menteri Pendidikan, Kebudayaan, Riset, dan Teknologi Nomor 5 Tahun 2022 tentang Standar Kompetensi Lulusan. (2024). Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi.
- Permendikdasmen No. 13 Tahun 2025 tentang Prinsip Pembelajaran Mendalam untuk Jenjang Pendidikan Anak Usia Dini, Pendidikan Dasar, dan Pendidikan Menengah. (2025). Kementerian Pendidikan Dasar dan Menengah.
- Pratiwi, R. P., Purnomo, M. E., Alwi, Z., & Zulkardi. (2023). Cultural representation in folklore the story of the three founding gods of the Besemah universe and its implications in literary learning. *English Review: Journal of English Education*, 11(1), 183–194. <https://doi.org/10.25134/erjee.v11i1.7263>
- Prihatmoko, A., Nurhayati, & Alwi, Z. (2025). A needs analysis for developing a flipbook-based e-module integrating Palembang puppet heritage in grade 12 Indonesian language learning. *Al-Ishlah: Jurnal Pendidikan*, 17(4). <https://doi.org/10.35445/alishlah.v17i4.7715>
- Purnomo, M. E., Ariska, M., Silvhiany, S., Dari, T. R. W., Nuzula, K., & Putri, A. Y. (2025). Pemulihan pembelajaran SMP region Palembang melalui pelatihan penguatan literasi dan numerasi. *Jurnal Pendidikan dan Pengabdian Masyarakat*, 8(1), 75–82. <https://doi.org/10.29303/jppm.v8i1.8458>
- Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). Free Press.
- Sholikhah, H. A., Prasetyo, C. T., Putri, A. Y., & Nuzula, K. (2026). Pelatihan guru dan pendampingan siswa SMAN Sumatra Selatan dalam penguatan keragaman dan kesantunan berbahasa melalui pemanfaatan media digital. *Jurnal Pendidikan dan Pengabdian Masyarakat*, 9(2). <https://doi.org/10.29303/jppm.v9i2.11413>