Volume 4 Number 1 (2026)
January – June 2026
Page: 141-156

Journal of Creative Power and Ambition (JCPA)
https://edujavare.com/index.php/jcpa
E-ISSN: 3031-4054
DOI: 10.70610/jcpa.v4i01.1147

Teacher AI Literacy in Digital Teaching Material Development: An Observational Study in Yogyakarta

Muhammad Fauzan Gustafi[1], Amalina Nur Arifah[2], Muhammad Haedar Zhafran Hidayatulloh[3]

[1,2,3]Universitas Siber Muhammadiyah, Indonesia;
* Correspondence e-mail; muhammadfauzangustafi@sibermu.ac.id

Article history Submitted: 2026/02/01; Revised: 2026/03/11; Accepted: 2026/04/23

Abstract

The rapid adoption of generative artificial intelligence in education has positioned Prompt Engineering as the primary AI interaction skill for educators. This study examines generative AI adoption patterns among educators in Yogyakarta, Indonesia, and investigates how engagement with prompt-based interfaces shapes adoption trajectories. Using a non-participant observational design, data were collected from 161 educators (49 early-career, 112 experienced) across four AI/AR/VR workshop sessions through structured field notes and purposive sampling across PAUD to higher education levels, then analyzed through reflexive thematic analysis. Results show that early-career educators demonstrated universal AI exposure (100%) compared to 72.32% among experienced educators, with cross-level analysis revealing that higher education lecturers showed notably higher senior-teacher exposure (93.3%) than primary (60.0%) and secondary (75.4%) levels. Three themes emerged: experience-based adoption differentials, prompt dependency as a cognitive bottleneck, and practice-based learning as the most effective adoption pathway. This study identifies prompt dependency as a transitional phenomenon in educator AI adoption, grounded in TAM, UTAUT, and Diffusion of Innovation frameworks, and proposes Context Engineering, the institutional-level design of AI interaction contexts, as the emerging competency educators need as language model instruction-following continues to improve. Findings contribute to technology adoption theory and provide actionable implications for AI literacy professional development policy in Indonesia.

Keywords AI Literacy, Digital Teaching Materials, Generative AI, Technology Adoption, Indonesian Teachers

INTRODUCTION

Generative artificial intelligence (generative AI) tools entered educational settings faster than most institutions could prepare for. Educators at every level now have access to systems capable of generating lesson plans, assessment items, and instructional feedback. Whether they are using these tools effectively is a separate question, and the answer depends heavily on who the educators are and what happens the first time they interact with AI (Yu & Guo, 2023; Harsanti et al., 2025). In Indonesia, national education reforms under Kurikulum Merdeka have explicitly positioned digital competency development as a priority, yet the gap between policy aspiration and actual classroom AI integration remains substantial (Raharjo & Rohmadi, 2024). As of 2025, Indonesian teachers' intention to adopt AI tools is moderated strongly by perceived complexity and institutional support, suggesting that the manner in which AI skills are introduced to educators carries significant consequences for adoption outcomes (Harsanti et al., 2025).

When generative AI first became widely accessible to non-technical users, Prompt Engineering (PE) was the primary skill educators were encouraged to learn. The logic was clear: structure your input correctly, using role assignments, few-shot examples, or chain-of-thought instructions, and the model would produce better outputs. Prior literature supports this directionally. White et al. (2023) catalogued reusable prompt patterns for ChatGPT across task categories, and Sahoo et al. (2024) surveyed PE techniques systematically across LLM deployments. In Indonesian education specifically, Gustafi (2025) applied structured PE to LMS content generation and demonstrated accuracy improvements, though with Mean Absolute Error (MAE) deviations that varied significantly by task type. These deviations indicate that individual prompt construction, even when carefully structured, remains variable in outcomes.

What received less attention during this period is how the underlying models changed. InstructGPT and its successors were trained with reinforcement learning from human feedback (RLHF) specifically to follow natural language instructions without requiring syntactically precise prompts (Ouyang et al., 2022). The gap between a carefully engineered prompt and a plainly worded request has narrowed considerably as model instruction-following has improved. Several studies have begun to document the implications: Zamfirescu-Pereira et al. (2023) found that non-expert LLM users routinely attribute output failures to prompt wording rather than to insufficient contextual information, cycling through rephrasing when the actual problem is that the model lacks the background to complete the task. Adopting TAM as an analytical frame, this behavior reflects a mismatch between users' mental model of AI interaction and the actual mechanisms that drive output quality, which in turn shapes perceived ease of use and adoption sustainability.

Against this background, the present study identifies a specific gap in the existing literature. Studies on educator AI adoption have examined adoption rates (Harsanti et al., 2025; Yang et al., 2025), AI literacy levels (Sari et al., 2025), and professional development design (Akanzire et al., 2024; Tan et al., 2024), but none have directly observed the interaction dynamics that emerge when educators engage with PE-based AI interfaces for the first time. The present study addresses this gap through structured observation of 161 educators across four workshop sessions in Yogyakarta, finding a pattern termed prompt dependency: educators newly introduced to PE techniques became absorbed in prompt mechanics rather than developing a working understanding of what contextual information the AI required. Unlike survey-based studies that rely on self-reported adoption intentions, the present study captures behavioral evidence from direct observation, which is the specific methodological contribution that differentiates it from prior work.

Three research questions frame the study. First, what are the generative AI adoption patterns among educators across educational levels in Yogyakarta? Second, what interaction dynamics emerge when educators first engage with prompt-based AI interfaces? Third, how does educators' focus on prompt mechanics affect sustainable AI adoption, and what does this imply for the concept of Context Engineering as a successor paradigm to PE? The argument this paper develops is that prompt dependency is a transitional phenomenon specific to the current phase of AI capability development. As models become better at interpreting plain-language instructions, the competency educators need is shifting from individual prompt crafting to what this paper calls Context Engineering: designing the AI interaction context at an instructional or institutional level, rather than constructing prompts case by case. This contributes to technology adoption theory by mapping a specific cognitive bottleneck within the TAM/UTAUT/Diffusion of Innovation framework and offers actionable implications for Indonesian AI literacy policy design.

Prompt Engineering is the practice of designing inputs to large language models to guide output quality, relevance, and structure. Gustafi (2025) operationalized PE within an LMS content generation task using six components: persona, context, task, format, exemplar, and tone. The six-component structure draws on and extends the prompt pattern catalog developed by White et al. (2023), which organized reusable prompt templates for ChatGPT across categories including input semantics, output customization, and error identification. Sahoo et al. (2024) surveyed PE techniques across major LLM deployments and identified zero-shot, few-shot, chain-of-thought, and self-consistency prompting as the most studied approaches. PE works better for some tasks than others: gains are most visible in multi-step reasoning and format-specific outputs. For routine content generation, the relationship between prompt complexity and output quality is nonlinear, and adding structural specificity past a certain point introduces new inconsistencies. The MAE variation across task types in Gustafi (2025) reflects this directly.

The Technology Acceptance Model (*TAM*) identifies perceived usefulness and perceived ease of use as the main determinants of individual technology adoption (Davis, 1989). In educational technology contexts, teachers' judgments about pedagogical relevance and accessibility drive sustained adoption more than institutional mandates (Xu et al., 2024). The Unified Theory of Acceptance and Use of Technology (*UTAUT*) extends this by adding performance expectancy, effort expectancy, social influence, and facilitating conditions as adoption predictors. Xu et al. (2024) applied UTAUT to Chinese university educators' adoption of AI tools and found that performance expectancy and social influence had the strongest effects on intention to use. Harsanti et al. (2025) applied a comparable framework to Indonesian teachers and found that institutional support and subjective norms moderated adoption intention, with perceived complexity as the main inhibiting factor.

Rogers' Diffusion of Innovation theory describes adoption as a population-level process, with individuals distributed across five categories: innovators, early adopters, early majority, late majority, and laggards (Rogers, 2003). Frei-Landau et al. (2022) applied this framework to mobile learning adoption among teachers during the COVID-19 period and found that generational and experiential differences corresponded to predictable differences in adopter category distribution. Yang et al. (2025) similarly identified experience-related asymmetries in AI literacy as a persistent structural barrier across a decade of research. The case for PE rested on an assumption that language models were sensitive to how instructions were phrased, not just what was being requested. For GPT-3 class models, this was largely accurate. The development of InstructGPT through RLHF changed this relationship: these models became better at inferring intent from plain-language instructions without requiring syntactically precise prompts (Ouyang et al., 2022). Zamfirescu-Pereira et al. (2023) documented the user-side consequence: non-experts attribute output failures to prompt wording rather than to insufficient context, producing iterative rephrasing rather than task-level problem analysis. The bottleneck in effective AI interaction is less about syntax and more about the user's capacity to identify and communicate the relevant instructional context. That is the gap Context Engineering addresses.

Figure 1 presents the integrated conceptual framework that guides the study's analysis. TAM and UTAUT explain individual-level adoption readiness, with perceived ease of use and effort expectancy identifying why early-career educators adopt AI more readily than experienced colleagues. Diffusion of Innovation theory explains the population-level distribution of these behaviors across adopter categories. The three frameworks converge at the interaction layer: when educators encounter PE techniques, effort expectancy increases if PE is perceived as complex (UTAUT), which activates the cognitive bottleneck this study terms prompt dependency. The output from this interaction determines whether educators develop context competency or remain fixated on prompt mechanics, with the former pathway leading toward the Context Engineering paradigm that this study positions as the target competency for sustainable AI adoption.

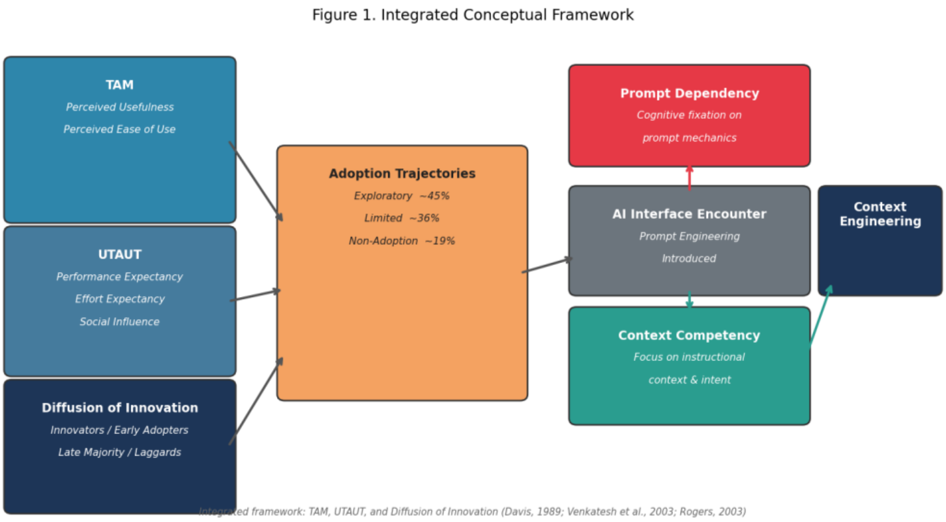

Figure 1. Integrated Conceptual Framework: TAM, UTAUT, and Diffusion of Innovation in Educator AI Adoption

METHODS

This study used a descriptive-qualitative design with non-participant structured observation as the primary data collection method. Non-participant observation was selected to minimize researcher influence on the natural interaction dynamics between educators and AI tools; observers recorded behaviors without actively facilitating or guiding the sessions. The research setting was four AI/AR/VR workshop sessions held in Yogyakarta, Indonesia, which are treated in this study as observation arenas rather than experimental interventions. Educators interacted with generative AI tools during practice activities, and observations were recorded as those interactions unfolded. The primary source of observational data is Gustafi et al. (2026), which documented field observations across these sessions.

Subjects were selected through purposive sampling. The inclusion criteria required that subjects were actively employed as educators at one of the following levels: PAUD, elementary, junior secondary, senior secondary/vocational, or higher education; that they were physically present and actively engaged throughout the observation sessions; and that they were attending the workshop through their institutional affiliation rather than as private individuals. A total of 161 educators met these criteria. Subjects were classified by teaching experience using a pragmatic operational threshold: early-career (young) teachers were those with fewer than five years of teaching experience (n = 49), and experienced (senior) teachers were those with five or more years (n = 112). This classification aligns with analogous distinctions in technology adoption research and is consistent with the experience thresholds used by Frei-Landau et al. (2022) in a comparable teacher technology adoption study.

All subjects participated voluntarily in the workshop sessions. Observation was conducted with the knowledge of session facilitators, and no identifying information was recorded in field notes. Consent was obtained through the workshop registration process, which informed participants that session activities might be documented for educational research purposes. The study met institutional review requirements through the workshop organizer's participant consent framework. Observers collected data through structured field notes, recording how educators engaged with AI tools, how they responded to PE instruction, their verbal and nonverbal reactions to AI-generated outputs, and the sequences of actions taken when outputs did not meet expectations. The observation guide was developed based on TAM and UTAUT constructs, mapping observable behaviors to theoretical indicators such as perceived ease of use (time spent on prompt adjustment), effort expectancy (expressions of difficulty or frustration), and performance expectancy (reactions to output quality). The observation guide was reviewed by two colleagues with expertise in qualitative educational research prior to deployment. No survey instruments or self-report measures were used.

Data analysis followed the reflexive thematic analysis approach described by Braun and Clarke (2006) and elaborated by Byrne (2022). The analytical process proceeded through six phases: familiarization with the data corpus through repeated reading of field notes; systematic generation of initial codes capturing observable behaviors and interaction patterns; aggregation of codes into candidate themes; review of candidate themes against the full dataset to verify adequate evidential grounding; definition and refinement of theme labels; and production of the final thematic report. Two researchers independently coded the full observation dataset;

coding divergences were resolved through discussion until consensus was reached.

Two validity strategies were employed. Triangulation was achieved by cross-referencing field observation notes with quantitative exposure data and with findings from Gustafi (2025), which provided independent evidence on PE-related behavior in a comparable educator population. Peer debriefing was conducted with a colleague external to the research team, who reviewed the emerging themes and challenged their interpretive basis. A potential limitation of the observational design is the Hawthorne effect: educators observed during workshop sessions may have modified their behavior due to awareness of observation. This risk was partially mitigated through the non-participant role of observers and the naturalistic workshop setting, which provided multiple simultaneous activities and reduced the salience of the observation function. Residual Hawthorne effects are acknowledged as a limitation in the interpretation of findings.

FINDINGS AND DISCUSSION

Overview of Research Subjects

The 161 educator subjects were distributed across educational levels as shown in Table 1. Secondary-level educators were the largest group (n = 79, 49.1%), followed by primary-level educators (n = 58, 36.0%) and higher education lecturers (n = 24, 14.9%). At every level, senior teachers substantially outnumbered early-career teachers. This distribution reflects the composition of the Indonesian educator workforce, where experienced teachers constitute the majority at all levels.

Table 1. Distribution of Research Subjects by Educational Level and Teaching Experience

Educational Level	Young Teachers (< 5 yrs)	Senior Teachers (> 5 yrs)	Total
Primary (PAUD/Elementary)	18	40	58
Secondary (Junior/Senior High)	22	57	79
Higher Education (Lecturer)	9	15	24
Total	**49**	**112**	**161**

Theme 1: Generative AI Adoption Patterns Across Experience Groups

Structured observation revealed substantial variation in prior AI exposure between early-career and experienced teachers. All 49 young teachers (100%) had prior exposure to generative AI tools before the workshop sessions, compared to 81 of 112 senior teachers (72.32%). The overall exposure rate across the full subject group was 80.75% (130 of 161 subjects). Table 2 summarizes these figures.

Table 2. Generative AI Exposure Rate by Teaching Experience Group

Teacher Group	n	AI Exposed	Not Exposed	Exposure Rate
Young Teachers	49	49	0	100.00%
Senior Teachers	112	81	31	72.32%
Total	**161**	**130**	**31**	**80.75%**

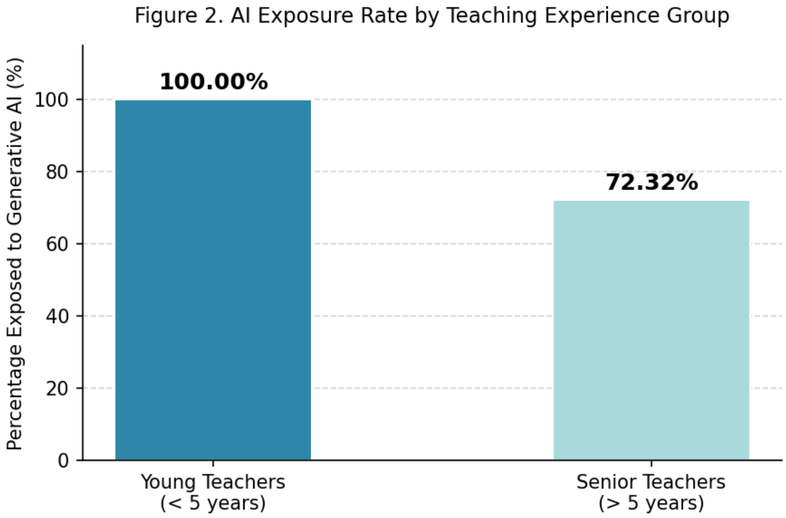

Figure 2. AI Exposure Rate by Teaching Experience Group

Cross-level analysis reveals a pattern that the two-group comparison alone conceals. Table 3 shows AI exposure rates disaggregated by both educational level and experience group. While the 100% exposure rate holds uniformly across all young teachers regardless of level, the experience-based gap differs substantially by level: primary school senior teachers show an exposure rate of 60.0%, compared to 75.4% among secondary teachers and 93.3% among higher education senior lecturers. The higher exposure rate among senior lecturers likely reflects the greater institutional digitalization of higher education contexts and the greater alignment between AI tool capabilities and the research and curriculum development tasks of university faculty.

Table 3. AI Exposure Rate by Educational Level and Teaching Experience Group

Educational Level	Young n	Young Exposed (%)	Senior n	Senior Exposed (%)	Total n	Total Exposed (%)
Primary (PAUD/Elementary)	18	100.0%	40	60.0%	58	72.4%
Secondary (Junior/Senior High)	22	100.0%	57	75.4%	79	82.3%
Higher Education (Lecturer)	9	100.0%	15	93.3%	24	95.8%
Total	**49**	**100.0%**	**112**	**72.3%**	**161**	**80.7%**

Within the exposed population, three adoption patterns were observed.

Exploratory adopters (approximately 45%) used AI tools independently to generate or modify teaching materials, though without a systematic pedagogical framework. Limited adopters (approximately 36%) were familiar with AI capabilities in concept but had not applied them to actual teaching content. Non-adopters (approximately 19%) had no prior generative AI engagement, whether due to access constraints, lack of confidence, or skepticism about the technology's relevance to their practice.

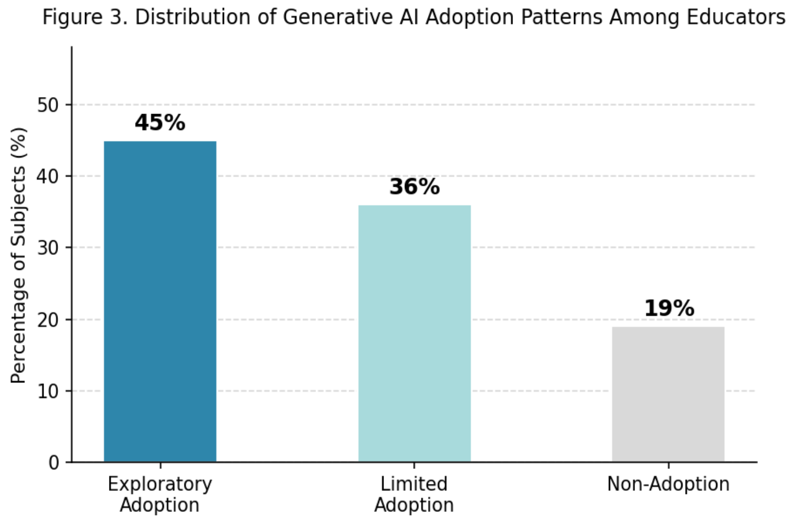

Figure 3. Distribution of Generative AI Adoption Patterns Among Educators

These patterns map onto Rogers' Diffusion of Innovation categories in expected ways. Young teachers, with universal exposure and higher rates of exploratory adoption, are the innovators and early adopters in this population. Senior teachers with prior exposure but limited applied use are the early or late majority. The non-adopter segment, almost entirely senior teachers, corresponds to the laggard category (Frei-Landau et al., 2022). Through a TAM lens, young teachers show higher perceived ease of use, which explains why exposure translates more consistently to active use in that group (Xu et al., 2024; Harsanti et al., 2025).

The cross-level data introduce one finding that runs against the simple experience-based prediction. Senior higher education lecturers show a notably higher exposure rate (93.3%) than their counterparts at primary (60.0%) and secondary (75.4%) levels, despite belonging to the same experience category. This suggests that institutional context and subject-matter characteristics moderate the experience-adoption relationship independently of the experience variable alone. Diffusion of Innovation adopter categories, which this study uses as a descriptive tool, may need to be applied with level-specific adjustments in future research on educator AI adoption. The national policy implication is that AI literacy interventions designed primarily for experienced teachers should be calibrated to institutional context rather than applied uniformly across levels.

Theme 2: Prompt Dependency and Its Consequences for AI Interaction

Among educators newly introduced to PE during the workshop sessions, a consistent behavioral pattern emerged. When guided through the six PE components (persona, context, task, format, exemplar, tone), this group redirected attention to the construction of the prompt itself rather than thinking through what the AI actually needed to know to help them. The following field note excerpts illustrate the pattern:

Field Note 01 (Session 2, Secondary School Teacher, Biology, 12 years' experience): The subject attempted to generate a lesson plan for a 10th-grade unit on ecosystem dynamics. Over approximately 14 minutes, the subject rewrote the prompt six times, changing phrases from 'explain about ecosystem' to 'generate a lesson about ecosystem topic' to 'create teaching material on ecosystem,' while keeping no additional context about grade level, learning objectives, or expected output format. After each attempt, the subject expressed dissatisfaction and asked the facilitator whether a specific formula existed for correct prompting. When the observer suggested specifying grade level, topic scope, and expected deliverable, the AI produced a structured lesson plan on the first subsequent attempt.

Field Note 02 (Session 3, Early-Career Mathematics Teacher, 2 years' experience): The subject generated a complete quiz set for fraction operations within approximately 4 minutes, without instruction on prompt structure. When asked about her approach afterward, she replied that she described what she needed as she would when asking a colleague. This naturalistic prompting behavior produced outputs comparable in quality to those generated by subjects who had received formal PE guidance.

These observations match the pattern documented by Zamfirescu-Pereira et al. (2023) in a controlled study of non-expert LLM users: when outputs were unsatisfactory, non-experts typically rephrased rather than reconsidered whether they had provided adequate context. The present study extends that finding to a naturalistic educational setting, where the behavioral consequence is measurable session time lost to prompt mechanics rather than task completion.

Independent evidence for the variability of PE outcomes comes from Gustafi (2025), who applied structured PE to LMS content generation tasks and measured output accuracy against target specifications. Figure 4 presents the MAE scores across four task types from that study.

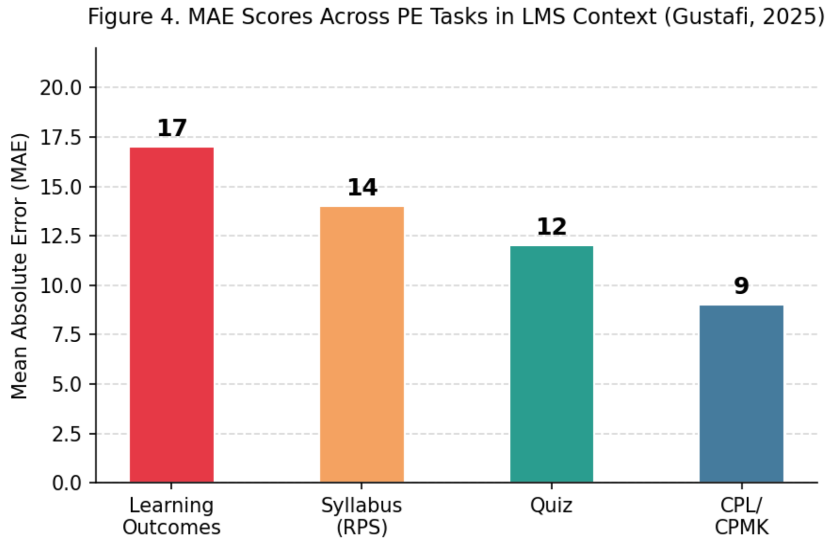

Figure 4. MAE Scores Across PE Tasks in LMS Context (Gustafi, 2025)

MAE values range from 9 for CPL/CPMK alignment tasks to 17 for learning outcome generation. Even carefully constructed prompts produced outputs with meaningful deviations from specifications in certain task types. For educators encountering this variability during initial AI exposure, the natural inference is that prompt refinement is the primary lever for quality improvement, which is exactly the framing that produces prompt dependency. The effort expectancy construct in UTAUT predicts that this perception will deter sustained adoption: if quality AI output is understood as requiring prompt expertise, educators who cannot identify the correct formula will likely retreat to non-adoption or limited use.

This cognitive fixation matters more now than it would have in 2022. RLHF-trained models respond to the intent expressed in plain language more than to syntactic structures (Ouyang et al., 2022), as illustrated clearly by Field Note 02: the early-career teacher who used naturalistic language achieved quality comparable to trained PE users. An educator whose mental model of AI interaction is primarily about prompt construction is therefore working with an increasingly outdated understanding of how these systems operate.

Theme 3: Practice-Based Learning as the Most Effective Adoption Pathway

Across all educator groups, the most consistent adoption catalyst was producing a usable output. Educators who shifted from skepticism or limited use to active adoption during the workshop sessions did so after generating a lesson plan draft, quiz set, or rubric that met or exceeded their quality expectations. Conceptual explanation alone did not move them.

This is consistent with what the literature on teacher professional development generally finds. Akanzire et al. (2024) found that hands-on exploration drove generative AI acceptance among teacher educators far more effectively than conceptual instruction. Rosmaria et al. (2024) reported the same pattern in an Indonesian elementary school context: the strongest adoption outcomes came when participants produced and evaluated complete instructional artifacts. Tan et al.'s (2024) systematic review of AI in teacher professional development reached the same conclusion across multiple contexts and levels.

The practical implication is that professional development programs should prioritize task completion over conceptual coverage. Raharjo and Rohmadi (2024) describe the Indonesian AI education context as one where awareness of AI capabilities has outrun actual integration into teaching workflows, a gap they attribute partly to training designs that emphasize tool introduction over practice-based use. Programs that keep participants in production mode, generating and evaluating real instructional content, are better positioned to build adoption sustainability than those that teach prompting as a standalone conceptual technique.

Overall Discussion: Teaching Experience, Institutional Context, and the Shift to Context Engineering

Teaching experience shapes the adoption trajectory in ways consistent across TAM, UTAUT, and Diffusion of Innovation frameworks. Younger educators adopt more rapidly because they enter with higher perceived ease of use and lower resistance to new tools. Experienced educators require stronger evidence of pedagogical value before consolidating adoption. Their standards are high because they have lived through many technology adoption cycles that promised transformation and delivered something more modest. The cross-level finding adds nuance: institutional context, not experience alone, determines where individual educators fall on the adoption curve. Senior lecturers at higher education institutions show near-universal AI exposure despite belonging to the experienced-teacher category, suggesting that the Diffusion of Innovation framework should be applied at the institutional context level, not just the individual experience level.

Prompt dependency, as this study argues, is a transitional phenomenon. PE was a legitimate and useful skill when it emerged, and its promotion as the primary AI literacy competency was contextually appropriate. The problem is that professional development frameworks have institutionalized PE-centric training just as model capabilities have reduced the marginal value of prompt precision. Teaching educators to optimize prompt syntax at this point in AI development may create habits that become liabilities as models continue to improve. This finding is not fully predicted by existing TAM, UTAUT, or Diffusion of Innovation models, which do not account for technology-side capability evolution as a moderator of the adoption-skill relationship. This is the specific theoretical gap the present study addresses.

The findings also have direct relevance to Indonesia's Merdeka Belajar national education agenda, which emphasizes student-centered learning and teacher creative autonomy. Generative AI offers genuine leverage for both objectives, but the prompt dependency pattern suggests a specific mechanism through which well-intentioned AI literacy training can produce unintended consequences. By centering PE as the primary competency, training programs risk creating fixation effects that impede the flexible, contextual AI use the curriculum envisions. Strengthening the linkage between AI literacy programs and Kurikulum Merdeka implementation frameworks is an actionable policy direction this study's findings support.

Figure 5 illustrates the paradigm shift this study argues is underway.

Figure 5. Paradigm Shift from Prompt Engineering to Context Engineering

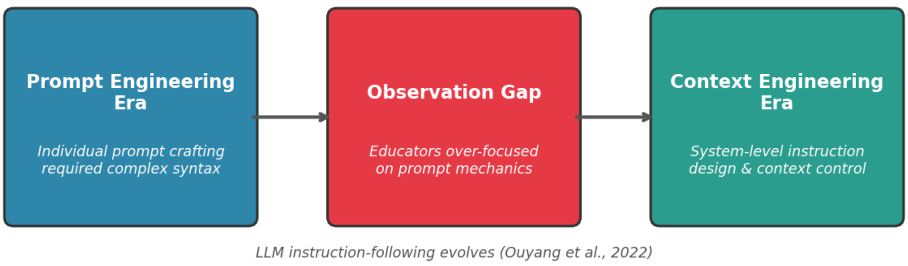

Figure 5. Paradigm Shift from Prompt Engineering to Context Engineering

Context Engineering, in contrast to PE, operates at the level of the interaction context rather than the individual prompt. Instead of training each educator to construct optimized prompts for every task, it involves pre-defining the instructional context, learner parameters, subject constraints, and output requirements that frame AI interactions across a course or institution. Enterprise AI deployments already work this way: system prompts, retrieval-augmented generation pipelines, and structured instruction templates are all forms of Context Engineering. Adapting these approaches for educational settings would reduce the cognitive demands on individual educators and produce more consistent outputs, without requiring each teacher to become proficient in prompt mechanics. Compared to studies in similar contexts (Yang et al., 2025; Sari et al., 2025; Xu et al., 2024), the present study is the first to identify prompt dependency as a direct observational finding and to propose the CE framework as its practical resolution, which constitutes the specific novel contribution of this work.

CONCLUSION

This study examined generative AI adoption patterns among 161 educators across PAUD to higher education levels in Yogyakarta through structured non-participant observation, finding that early-career teachers showed universal AI exposure (100%) compared to 72.32% among experienced teachers, with cross-level analysis revealing that institutional context moderates this gap: senior higher education lecturers showed a 93.3% exposure rate versus 60.0% and 75.4% at primary and secondary levels respectively. Three themes emerged from reflexive thematic analysis: experience-based adoption differentials consistent with TAM, UTAUT, and Diffusion of Innovation predictions; prompt dependency as a cognitive bottleneck in which educators newly introduced to PE techniques become fixated on prompt mechanics rather than instructional context, directly supported by observational field notes and corroborated by PE output variability data from Gustafi (2025); and practice-based task completion as the most effective adoption catalyst, consistent with the broader professional development literature. The theoretical contribution of this study is the identification of prompt dependency as a transitional phenomenon produced by the misalignment between PE-centric training frameworks and the rapidly improving instruction-following capabilities of contemporary language models, a dynamic not captured by existing TAM, UTAUT, or Diffusion of Innovation models. For Indonesian AI literacy policy, the study recommends two specific actions: professional development programs should shift from PE instruction toward context design workshops that require educators to specify learning contexts, learner profiles, and pedagogical objectives before any AI interaction; and national AI literacy frameworks should incorporate Context Engineering as a target competency, with assessment criteria focused on instructional context specification rather than prompt syntax. Limitations include the single-city geographic scope, the five-year experience threshold as an operational rather than theoretical construct, the Hawthorne effect risk inherent in observed workshop settings, and the absence of interview data to validate observational inferences. Future research should develop and evaluate Context Engineering frameworks for educational settings and conduct longitudinal studies to determine whether prompt dependency resolves independently over time or requires targeted intervention.

REFERENCES

Akanzire, B. N., Nyaaba, M., & Nabang, M. (2024). Generative AI in teacher education: Teacher educators' perception and preparedness. Journal of Digital

Educational Technology, 5(1), ep2508. https://doi.org/10.30935/jdet/15887

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Byrne, D. (2022). A worked example of Braun and Clarke's approach to reflexive thematic analysis. Quality & Quantity, 56(3), 1391–1412. https://doi.org/10.1007/s11135-021-01182-y

Frei-Landau, R., Muchnik-Rozanov, Y., & Avidov-Ungar, O. (2022). Using Rogers' diffusion of innovation theory to conceptualize the mobile-learning adoption process in teacher education in the COVID-19 era. Education and Information Technologies, 27(9), 12811–12838. https://doi.org/10.1007/s10639-022-11148-8

Gustafi, M. F. (2025). Optimalisasi penggunaan AI dalam bahan ajar penerapan teknik prompt engineering untuk learning management systems (LMS). Jurnal Riset Sains dan Teknologi, 9(2), 121–133. https://doi.org/10.30595/jrst.v9i2.23393

Gustafi, M. F., Arifah, A. N., & Hidayatulloh, M. H. Z. (2026). Workshop literasi AI untuk pembuatan bahan ajar digital bagi guru SD hingga SMA/SMK. Jurnal Penelitian dan Pengabdian Masyarakat, 4(1), 47–55. https://doi.org/10.61231/evcn1241

Harsanti, H. R., Sudibjo, N., Riady, S., & Yu, P. (2025). Exploring Indonesian teachers' intention to use artificial intelligence in schools. Cogent Social Sciences, 11(1). https://doi.org/10.1080/23311886.2025.2593596

Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A., Schulman, J., Hilton, J., Kelton, F., Miller, L., Simens, M., Askell, A., Welinder, P., Christiano, P., Leike, J., & Lowe, R. (2022). Training language models to follow instructions with human feedback. Advances in Neural Information Processing Systems, 35, 27730–27744.

Raharjo, R. S., & Rohmadi, S. H. (2024). Artificial intelligence in Indonesian education: A critical review of ethical considerations, implementation challenges, and educational management perspectives. At-Tarbawi: Jurnal Kajian Kependidikan Islam. https://ejournal.uinsaid.ac.id/at-tarbawi/article/view/12141

Rosmaria, Utama, H. B., Sari, R. E., Pratiwi, R., Sepkia, Y., Masnur, A., Stevani, A. F., Andrean, J. F., & Adli, M. (2024). Pelatihan pengembangan bahan ajar digital berbasis artificial intelligence (AI) bagi guru SD Gugus III Kecamatan Baso. J-COSCIS: Journal of Computer Science Community Service.

https://journal.unilak.ac.id/index.php/jcoscis/article/view/25860

Sahoo, P., Singh, A. K., Saha, S., Jain, V., Mondal, S., & Chadha, A. (2024). A systematic survey of prompt engineering in large language models: Techniques and applications. arXiv:2402.07927. https://arxiv.org/abs/2402.07927

Sari, D. K., Supahar, S., Rosana, D., Dinata, P. A. C., & Istiqlal, M. (2025). Measuring artificial intelligence literacy: The perspective of Indonesian higher education students. Journal of Pedagogical Research, 9(2), 143–157.

Tan, X., Cheng, G., & Ling, M. H. (2024). Artificial intelligence in teaching and teacher professional development: A systematic review. Computers & Education: Artificial Intelligence. https://www.sciencedirect.com/science/article/pii/S2666920X24001589

White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., Elnashar, A., Spencer-Smith, J., & Schmidt, D. C. (2023). A prompt pattern catalog to enhance prompt engineering with ChatGPT. arXiv:2302.11382. https://arxiv.org/abs/2302.11382

Xu, S., Chen, P., & Zhang, G. (2024). Exploring Chinese university educators' acceptance and intention to use AI tools: An application of the UTAUT2 model. SAGE Open, 14(4). https://doi.org/10.1177/21582440241290013

Yang, Y., Zhang, Y., Sun, D., He, W., & Wei, Y. (2025). Navigating the landscape of AI literacy education: Insights from a decade of research (2014–2024). Humanities and Social Sciences Communications, 12, Article 583. https://doi.org/10.1038/s41599-025-04583-8

Yu, H., & Guo, Y. (2023). Generative artificial intelligence empowers educational reform: Current status, issues, and prospects. Frontiers in Education, 8, Article 1183162. https://doi.org/10.3389/feduc.2023.1183162

Zamfirescu-Pereira, J. D., Wong, R. Y., Hartmann, B., & Yang, Q. (2023). Why Johnny can't prompt: How non-AI experts struggle to prompt LLMs. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23). ACM. https://doi.org/10.1145/3544548.3581388