International Journal of Learning, Teaching and Educational Research
Vol. 25, No. 7, pp. 379-403, July 2026
https://doi.org/10.26803/ijlter.25.7.17
Received Feb 6, 2026; Revised Apr 30, 2026; Accepted Jun 25, 2026

Language and Communication Students' Perceived Mastery in AI Chatbot Prompt Engineering: A Study on Vibe Coding in Educational Mobile Application Development

Nur Izzati Khairuddin
INTI International University
Negeri Sembilan, Malaysia

Muhammad Haziq Abd Rashid*, **Hairul Azhar Mohamad**, **Amir Lukman Abd Rahman**, **Pavithran Ravinthra Nath**
Akademi Pengajian Bahasa, Universiti Teknologi MARA
Shah Alam, Selangor, Malaysia

Urai Salam
Universitas Tanjungpura Kalimantan
Pontianak, Indonesia

Maksetbay Mambetniyazov
University of Innovation Technologies
Republic of Karakalpakstan, Uzbekistan

Abstract. This study examines language and communication students' self-perceived mastery of AI chatbot prompt engineering, with particular attention to "vibe coding" for educational mobile application development. In this study, vibe coding refers to the deliberate use of prompts to shape tone, style, audience orientation, and user-facing educational content. Using a pattern-based framework centred on roles, constraints, and examples, the study investigates how students combine structure and exploratory prompting. A quantitative cross-sectional survey was administered online during March–July 2025 (Semester 2). Complete responses from 120 undergraduates were analysed from approximately 126 invited students (analytic response rate = 95.2%). The questionnaire demonstrated excellent overall internal consistency (Cronbach's α = .965). Results indicate that students usually begin tasks with structured prompts but later move towards mixed or unstructured prompting styles, suggesting a control-then-explore sequence. Longer

Citation:
Khairuddin, N. I., Rashid, M. H. A., Mohamad, H. A., Rahman, A. L. A., Nath, P. R., Salam, U., & Mambetniyazov, M. (2026). Language and Communication Students' Perceived Mastery in AI Chatbot Prompt Engineering: A Study on Vibe Coding in Educational Mobile Application Development. *International Journal of Learning, Teaching and Educational Research*, *25*(7), 379–403.
https://doi.org/10.26803/ijlter.25.7.17

*Corresponding author: Muhammad Haziq Abd Rashid; *haziqrashid@uitm.edu.my*

exposure to AI chatbots and more extensive prompt-engineering training were associated with higher self-perceived competency and output efficiency, whereas weekly usage frequency did not show statistically significant differences. Educational background was associated with prompting style and usage, but not with overall perceived competency or output efficiency, and gender differences were negligible. The findings suggest that scaffolded instruction in prompt engineering can support more confident and consistent student use of AI chatbots. Future research should triangulate self-reports with behavioural logs, archived prompts, and performance-based outputs.

Keywords: prompt engineering; AI chatbot; vibe coding; mobile application development; sustainable development education

1. Introduction

The choice between structured and unstructured prompt engineering significantly influences the dependability of language models used in educational app development. Structured prompts establish clear roles, objectives, formats, and provide examples, offering a framework for predictable outcomes, whereas unstructured prompts allow for more open-ended exploration. Research involving non-experts in AI reveals that creating effective prompts does not come naturally, and approaches relying solely on instructions tend to produce fragile results (Zamfirescu-Pereira et al., 2023).

In real-world settings, best practices recommend using explicit instructions, defined boundaries, and example-based formats to enhance consistency and quality (OpenAI, 2025). Although unstructured prompts may encourage creative ideas, they often cause greater variability and require repeated revisions when precision and adherence to specific formats are essential. Structured prompting serves as a support system for students as they develop mental models of chatbot reasoning. This suggests that teaching methods should prioritise pattern-based examples while maintaining opportunities for exploratory experimentation.

Demographic and experiential factors influence how quickly students develop confidence in prompt engineering, but their effects are inconsistent. Research indicates that previous experiences with AI education and the context in which learning occurs have a stronger impact on AI literacy than factors like age or academic year (O'Dea et al., 2024). Beyond these factors, programme culture, assessment styles, and exposure to prompt-related activities also play important roles. For instance, students in technically focused programmes tend to adopt structured prompting earlier, whereas those in language-centred disciplines initially rely more on open-ended phrasing before settling into set templates. Furthermore, while frequent weekly use helps with familiarity, active engagement and specialised workshops are more effective at consolidating skills.

Therefore, demographic categories hold less weight compared to individual learning experiences, indicating that providing equitable access to guided practice and clear instruction is key to reducing differences in perceived mastery over time. Together, these insights reveal that structured prompts provide a reliable foundation, while unstructured approaches encourage creative exploration. Students achieve the greatest benefit when teaching blends both methods. Since hands-on experience and formal training carry more influence than fixed demographic characteristics, educational programmes offering equitable and scaffolded learning opportunities can enhance perceived mastery for all students without restricting the creative latitude that makes vibe coding distinctive.

In this study, AI literacy refers to students' self-reported understanding of how AI chatbots function and how they may be used appropriately in academic work; prompt engineering competency refers to their perceived ability to formulate clear, purposeful prompts using elements such as roles, constraints, examples, and output guidance; vibe-coding prompting style and usage refers to the ways students structure prompts to shape tone, voice, audience alignment, and interactional feel when developing educational mobile-app content; and perceived output efficiency refers to students' judgement that a prompting approach helps them obtain usable responses with less trial and error. Importantly, the present study investigates perceived mastery rather than objective prompt-engineering performance (Knoth et al., 2024; Walter, 2024; White et al., 2023).

1.1 Problem Statement

Although prompt engineering has become widely acknowledged as an essential skill for making the most of large language models, its formal incorporation into language and communication curricula remains sparse. Recent syntheses show that higher-education work on Generative AI is evolving from conceptual mapping to applied classroom interventions, but prompt-engineering pedagogy and tone/affect ("vibe") control in student-authored prompts remain under-specified, especially in mobile learning contexts (Yang et al., 2024; Borromeo et al., 2025). Many students engage with AI tools without clear strategies, which often leads to inconsistent results and inefficient workflows.

Recent research highlights that prompt engineering involves iterative design, critical thinking, and subject-matter knowledge, yet these skills are rarely taught explicitly in higher education settings (Cain, 2024). Similarly, beginners often rely on trial and error, prolonging their learning process and diminishing their confidence in using AI effectively (Giray, 2023). This situation reveals a gap: while AI tools are easily accessible, mastery of structured prompting techniques is far from assured. Therefore, there is an urgent need for targeted teaching methods that connect theoretical knowledge with practical application.

Producing complex and stylistically sophisticated output remains a challenge for AI chatbots, especially when applied to educational apps. Although AI tools generally improve grammar and vocabulary, they often fall short in maintaining stylistic consistency and rhetorical suitability in complex language tasks (Zhao,

2024). These chatbots also frequently struggle with understanding context and interpreting nuanced intentions, which compromises their ability to produce responses that accurately reflect user goals (Fritsch, 2024). For students specialising in language and communication, these shortcomings restrict their capacity to develop authentic content tailored to specific audiences. As a result, students need to learn how to design prompts that steer AI towards the precise linguistic and stylistic results they seek, emphasising prompt engineering as both a technical craft and a rhetorical skill.

Demographic factors such as previous exposure to AI, academic discipline, and gender also play a role in shaping students' AI literacy and confidence with prompt engineering. Studies indicate that students from technical backgrounds or with AI-related coursework generally demonstrate greater proficiency in AI-related tasks than their peers from non-technical areas (Toker Gokce et al., 2024). Frequent AI-tool use and training opportunities also strongly correlate with higher competency levels, although gender-based differences in technical understanding persist. These disparities suggest that without fair and structured training opportunities, such demographic differences may continue to contribute to unequal mastery levels, hindering inclusive adoption of AI technologies in educational settings.

Taken together, these factors point to a pressing research gap. Limited structured instruction in prompt engineering, difficulties in obtaining stylistically precise AI output, and demographic disparities collectively underline the need for a better understanding of students' perceived mastery in this area. Although AI chatbots are increasingly common in educational contexts, little is known about how language and communication students perceive their skills in prompt engineering, particularly concerning vibe-coding tasks in mobile app development. This study seeks to fill that gap by investigating perceived mastery, prompting preferences, and demographic influences to inform effective teaching approaches for AI literacy.

The primary aim of this research is to explore how language and communication students perceive their mastery of AI chatbot prompt engineering, focusing especially on their prompting styles and the factors that influence them within the context of vibe-coding tasks for educational mobile app development. This inquiry intends to uncover patterns, relationships, and demographic effects shaping students' abilities to craft effective ChatGPT prompts.

The research questions guiding this study are:

1. What are the common ChatGPT prompt structures employed by language and communication students when working on vibe-coding tasks?
2. How does exposure to AI chatbot instruction affect students' prompting styles?
3. To what extent does academic background influence students' approaches to ChatGPT prompting?
4. Does gender have any impact on students' ChatGPT prompting styles?

2. Literature Review

The development of prompt engineering has shifted from informal tips to organised pattern libraries, with White et al. (2023) playing a key role by introducing a pattern-based framework that captures reusable strategies for structuring prompts and combining them across different tasks. This framework provides a catalogue of "prompt patterns," a schema for describing them, and guidance on how these patterns can be combined to guide language models towards consistent, reliable outputs, avoiding unstable or unpredictable responses. For vibe-coding tasks in educational app development, these patterns are applied by clearly defining roles, setting constraints, and using examples to control tone, format, and audience.

Previous scholars have argued that novices benefit more from structured guidance than from random trial and error (Zamfirescu-Pereira et al., 2023) and have called for explicit teaching of prompt frameworks in higher education (Lee & Palmer, 2025). The framework itself offers a practical tool to deliver such guidance, reframing prompting as a deliberate design process based on reusable forms while establishing a shared vocabulary that helps students understand why certain prompts succeed.

Linking this framework to prior research and the current study highlights its relevance. Across the educational field, scholars note the potential of large language models to assist with content creation and interaction but also stress that benefits depend heavily on users' literacies and disciplined prompting, issues that the White et al. framework directly addresses (Kasneci et al., 2023). Reviews of chatbot use in higher education describe the field as fragmented and lacking shared conceptual anchors, emphasising the need for pattern-based methods to promote consistent practice (McGrath et al., 2024).

In language and communication studies, students must balance style, voice, and structure when co-creating with AI; here, combining prompt patterns helps set register, genre, and rhetorical direction to preserve authorial intent (Wang, 2024). Applied to vibe coding, this framework offers clear, teachable steps, such as defining roles, specifying expected outputs, and including examples, that can enhance perceived mastery. Hence, it provides a practical scaffold for analysing prompting styles and their influences in mobile app development projects.

The discussion around White et al.'s (2023) framework presents a solid theoretical foundation for seeing structured prompting as an intentional design activity rather than a hit-and-miss approach. Studies affirm that structured prompts improve clarity and output quality, though adapting these patterns to creative and context-sensitive tasks like vibe coding remains challenging. Furthermore, despite its practical value, empirical investigations into the framework's application among language and communication students, especially in mobile app development, are scarce. These research gaps highlight the need for studies that not only verify the framework's usefulness but also explore how demographic and experiential factors influence its effectiveness in real educational settings.

Recent scholarship further suggests that AI literacy and prompt engineering should be treated as related but distinct educational capabilities. Knoth et al. (2024) show that students' AI literacy helps explain differences in prompting behaviour and output quality, while Walter (2024) argues that AI literacy, prompt engineering, and critical thinking should be integrated together in higher-education practice. These perspectives strengthen the present study's distinction between AI literacy, prompt engineering competency, prompting style and usage, and perceived output efficiency.

2.1 Common ChatGPT Prompt Structures among Language and Communication Students

When examining common ChatGPT prompt structures among language and communication students, research finds that learners often start with vague instructions before gradually incorporating examples and role definitions as tasks become more complex. For instance, a lab study involving ten non-experts found that participants approached prompt editing in an ad hoc manner, which limited output consistency (Zamfirescu-Pereira et al., 2023). Similarly, a phenomenological study with six first-year writing students revealed that while ChatGPT aided idea generation and organisation, students struggled to maintain consistent voice and coherence, leading them to introduce clearer formatting cues (Wang, 2024). These findings suggest that students naturally gravitate towards structured elements like exemplars, roles, and constraints after initial exploration, pointing to the need for scaffolds normalising these features.

Additional evidence from a two-week field intervention with 11 language teachers showed that prompts specifying roles, content boundaries, and evaluative tasks helped control genre and register, although teachers moderated chatbot outputs to align with pedagogical goals (Jeon & Lee, 2023). A review of 23 post-ChatGPT studies in higher education noted the absence of shared prompting taxonomies, complicating attempts to establish standardised practices (McGrath et al., 2024). Evidence from undergraduates indicates that perceived usefulness, ease of use, and social influence are reliable predictors of AI adoption; thus, perceived mastery in prompt engineering should improve when task scaffolds and examples are clear and socially endorsed (Kowang et al., 2025; Wong et al., 2025). Thus, even as structured prompting supports style and format consistency, programmes must offer common templates and reflective routines to help students adapt patterns to diverse rhetorical contexts.

2.2 Influence of AI Chatbot Lessons on Prompt Engineering Mastery in Vibe Coding

Regarding the impact of AI chatbot lessons on prompt engineering mastery in vibe coding, large-scale cross-sectional data confirm the importance of instruction. For example, a decision-tree analysis of 664 undergraduates found that completing AI-related courses predicted stronger AI literacy, especially when coupled with regular AI tool use (Toker Gokce et al., 2024). Similarly, a survey of 234 students from the UK and Hong Kong demonstrated that prior AI education significantly boosted generative AI literacy across different settings (O'Dea et al., 2024). Complementing these findings, a systematic review recommends embedding framework-based prompting within curricula to elevate output

quality and facilitate transfer of skills (Lee & Palmer, 2025). Targeted lessons therefore appear to accelerate mastery by making prompt patterns explicit and measurable.

Short workshops combined with guided practice also improve adoption and self-confidence. For instance, a UTAUT-based survey of 74 undergraduates found that perceived productivity benefits and mentor guidance influenced chatbot use more than demographic factors, suggesting instruction can shape behaviours (Patterson et al., 2024). Likewise, a two-week professional development study with 11 teachers reported that role and constraint prompt enhanced alignment with tasks but required careful pedagogical support to avoid over-dependence (Jeon & Lee, 2023). Commentaries stress the need to build competency due to ongoing issues with the brittleness of large language models (Kasneci et al., 2023). Therefore, lesson designs should blend pattern drills with opportunities for reflection to prevent superficial prompting.

2.3 Impact of Educational Backgrounds on Prompt Engineering Mastery in Vibe Coding

When considering the impact of educational backgrounds on prompt engineering mastery, multiple studies reveal differences related to discipline. For example, among 664 students, those in engineering, technology, and natural sciences displayed higher AI literacy compared to peers in education and sports sciences, signalling the influence of programme culture on practice (Toker Gokce et al., 2024). In another survey of 234 students across sites, prior learning and institutional context predicted literacy more than age or academic level, with some variation across countries (O'Dea et al., 2024). A higher education review confirms uneven conceptual frameworks across disciplines (McGrath et al., 2024). This suggests curricula should adapt prompting patterns to fit disciplinary genres, enabling students to transfer structured approaches to domain-specific tasks.

Further adoption research indicates that background interacts with use patterns. The UTAUT survey (n=74) linked chatbot use more strongly to perceived productivity and task variety than to fixed traits, implying that advantages from background can be offset by well-designed tasks (Patterson et al., 2024). In teacher professional development settings, disciplinary goals shaped the balance of role-setting and evaluative prompts in lesson planning (Jeon & Lee, 2023). Commentators highlight that literacy development is the first step to prompting mastery, equipping various benefits in the learning curve (Kasneci et al., 2023). Therefore, while disciplinary training offers a head start, structured practice and feedback mechanisms can equalise mastery in vibe-coding workflows.

2.4 Influence of Gender on Prompt Engineering Mastery in Vibe Coding

Gender's influence on prompt engineering mastery emerges as something complex. In a sample of 664 students, males scored somewhat higher on technical understanding, though differences varied across specific skills and usage levels (Toker Gokce et al., 2024). Conversely, a UTAUT survey of 74 undergraduates found no significant gender effect on chatbot use, with productivity beliefs playing a more crucial role (Patterson et al., 2024). Research among secondary

students (n=622) showed girls had lower attitudes towards AI learning and literacy, but the link between attitude, literacy, and career interest was similar for both genders (Zhang et al., 2025). The findings suggest that observed gender differences may stem from disparities in confidence and exposure rather than ability, indicating the value of targeted support.

Training conditions appear to reduce gender gaps. The 664-student decision-tree study revealed that enrolment in AI courses and frequent AI tool use helped moderate group differences, implying that programme design can reduce disparities (Toker Gokce et al., 2024). Similarly, findings from 74 undergraduates emphasise mentoring and task diversity as stronger predictors of continued chatbot use than demographic factors (Patterson et al., 2024). A sector-wide review advocates fostering shared literacies to promote consistent practices across mixed cohorts (McGrath et al., 2024). In light of this, inclusive teaching strategies that combine pattern catalogues, guided practice, and low-pressure exercises can empower all students to build confidence with structured prompting in vibe coding, regardless of gender.

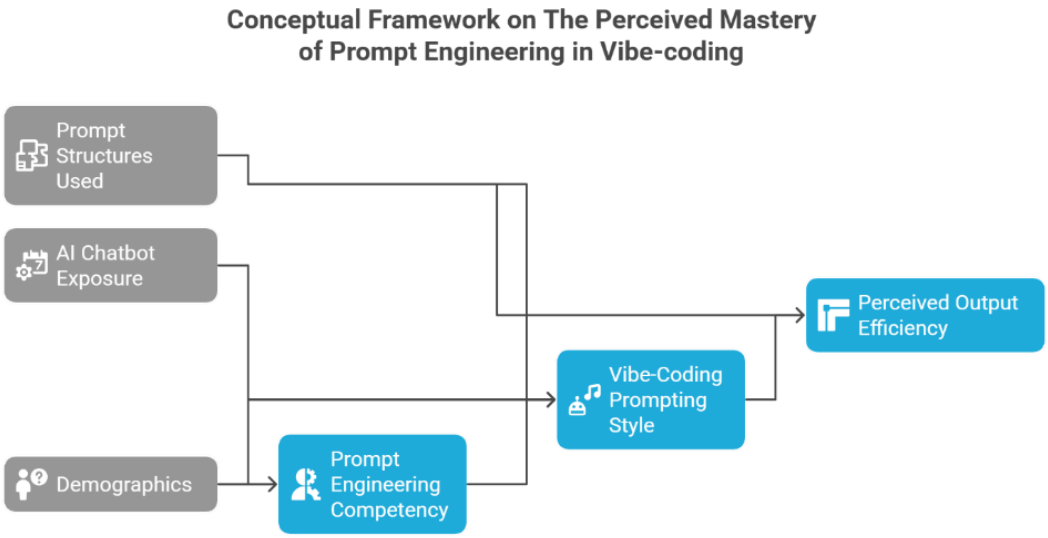

Figure 1: The conceptual framework on the perceived mastery of prompt engineering in vibe-coding

These studies converge on clear patterns: structured prompts enhance clarity and reliability, while unstructured prompts foster creativity but may sacrifice consistency. Instructional interventions, whether brief workshops or full courses, are crucial for improving mastery, yet their integration into curricula is patchy. Background and gender exert nuanced influences, often mediated by prior experience and confidence, indicating that demographic gaps are neither fixed nor insurmountable. However, most existing research relies on self-reported data or small samples, revealing a need for studies employing robust quantitative methods combined with real-world contextual tasks such as vibe coding. This would better clarify how structured training and demographic variables jointly affect perceived mastery.

Figure 1 shows the conceptual framework on the perceived mastery of prompt engineering in vibe-coding. Based on the pattern-based approach to prompting and recent empirical findings, the framework views AI exposure and demographic factors as external influences that shape students' perceived mastery and the prompting styles they employ in vibe coding. It differentiates between prompt structures; structured, unstructured, or mixed, and the actual prompting styles and behaviours reported by students as they engage with tasks.

The model acknowledges that learners typically start with flexible, less organised prompts but tend to adopt more structured approaches as task complexity increases. It also considers perceived output efficiency as a result influenced by both skill level and style, recognising that clear roles, constraints, and examples help stabilise tone, format, and audience appropriateness. The diagram therefore maps the pathway from students' background and experience through to mastery, prompting style, and perceived efficiency.

First, AI chatbot exposure, measured by cumulative months of use and participation in focused lessons is expected to enhance perceived competence, encourage more disciplined prompting styles, and boost perceived efficiency. Second, demographic variables such as educational background and gender may influence students' initial approaches and confidence, although these effects often depend on access to instruction. Third, the choice of prompt structures, whether structured or mixed, should guide students towards styles that better manage register and format, thereby increasing efficiency.

Finally, the relationship between competency, style, and efficiency reflects the practical sequence seen in vibe coding: stronger competencies lead to more purposeful prompting styles, which in turn support outputs that students regard as more efficient. Relating directly to the study's objectives, this framework clarifies how AI exposure, demographic characteristics, and choices in prompt structuring collectively predict students' perceived mastery in prompt engineering. Moreover, it offers testable pathways that can guide the development of effective course designs for teaching vibe coding in educational mobile app development.

3. Methodology

3.1 Research Design

This study employs a quantitative, cross-sectional research design to investigate the relationships among students' AI chatbot exposure, demographic characteristics, and their self-perceived mastery of prompt engineering in vibe coding. Quantitative methods are common in education research because they allow variables to be measured numerically and compared statistically across groups (Siripipatthanakul et al., 2023). In the present study, the design is intended to identify patterns of association rather than to support causal inference. By combining a structured questionnaire with inferential statistics, the study examines prompting styles and competency perceptions within an authentic educational setting.

The study relies on a questionnaire-based approach paired with inferential statistical analysis to examine associations among the focal variables. Structured questionnaires are useful for generating standardised responses from relatively large student cohorts, thereby supporting comparison across groups and transparent statistical testing (Yusoff et al., 2021). The instrument was adapted from prompt-pattern studies and aligned with the study constructs of AI literacy, prompt engineering competency, vibe-coding prompting style and usage, and perceived output efficiency.

3.2 Sample Size and Sample Type

The analysed sample comprised 120 undergraduate students from a Malaysian public university enrolled in language and communication courses. Approximately 126 students were invited to participate, yielding 120 complete responses for analysis (analytic response rate = 95.2%). This sample size is comparable to prior studies involving student interaction with generative AI in educational contexts (Wang, 2024; Woo et al., 2024). Selecting language and communication students suits the study's focus on prompting styles, where linguistic precision, register, and rhetorical appropriateness are central.

3.3 Sampling Method

A purposive sampling approach was adopted to recruit students majoring in English and enrolled in relevant e-language or language-and-communication courses during March – July 2025 semester. All students in the identified classes were invited to participate through course communication channels, but only complete submissions were retained for analysis. Purposive sampling is appropriate when the research requires participants with characteristics directly linked to the study aims (Nyimbili & Nyimbili, 2024). In this case, the approach ensured that respondents had sufficient engagement with language-intensive tasks and AI chatbot use to comment meaningfully on prompt engineering in vibe coding.

3.4 Instrument to Collect the Data

Data were gathered via a Likert-scale questionnaire adapted from White et al.'s (2023) prompt engineering framework and organised into five parts: demographics, AI literacy, prompt engineering competency, vibe-coding prompting styles and usage, and perceived output efficiency. Before the main administration, the questionnaire underwent small-scale pilot testing and internal expert review to improve item clarity, content alignment, and feasibility, which is consistent with established survey-development practice (Bujang et al., 2024; Yusoff et al., 2021). The analysed instrument demonstrated excellent overall internal consistency (Cronbach's $\alpha = .965$) across 46 scored items.

3.5 Data Collection Procedure

The questionnaire was administered online from March to July 2025 (Semester 2) to support ease of access and response efficiency. Participants were undergraduate students enrolled in relevant e-language and language-and-communication courses. Approximately 126 students were invited, and 120 complete responses were retained for analysis. Only complete submissions were included in the final dataset. Responses were stored securely in a password-

protected digital repository before cleaning and analysis. This procedure is consistent with common survey administration practices in educational research (Rokeman, 2024; Woo et al., 2024).

3.6 Data Analysis Procedure

Analysis was conducted using SPSS. Descriptive statistics summarised participant characteristics and central tendencies, while independent-samples t-tests examined gender differences and one-way ANOVAs compared differences by educational background, months of AI chatbot use, weekly usage frequency, and prior prompt-engineering training. Effect sizes were reported alongside inferential tests to improve interpretation. Because the study analysed multi-item composite scores rather than single Likert items, parametric tests were used as an approximation for interval-level measurement; however, this choice should be interpreted cautiously and transparently, especially in questionnaire research (Rokeman, 2024).

3.7 Validity and Reliability Tools

To verify instrument quality, the questionnaire underwent small-scale pilot testing and internal expert review before the main study. Internal consistency reliability for the analysed questionnaire was excellent, with an overall Cronbach's alpha of .965 across 46 scored items. Cronbach's alpha remains a widely used indicator of internal consistency in educational research, although it should be interpreted alongside evidence of content alignment and instrument development procedures (Ahmad et al., 2024; Zakariya, 2022). In the present study, the pilot test and expert review supported content validity by refining wording, clarity, and construct alignment prior to full administration.

4. Results and Findings

Table 1 presents the distribution of prompt structures used at the beginning of tasks, while Table 2 summarises prompting styles reported over the preceding month. At task onset, the majority of students favoured structured prompts (78.3%, n = 94), with far fewer starting with unstructured prompts (21.7%, n = 26). This suggests a clear preference for establishing constraints and roles upfront. However, when looking at prompting styles used in the last month, the distribution was more evenly spread; structured prompts accounted for 34.2% (n = 41), unstructured 33.3% (n = 40), and mixed prompts 32.5% (n = 39). This balance points to a more adaptable, day-to-day approach.

These patterns suggest that students generally begin tasks with structured, controlled prompts to set clarity but then shift toward mixed or unstructured styles as the work progresses, inviting iteration and creative adjustment. In practical terms, starting with a structured prompt helps establish clear direction, while later switching prompts supports exploration and refinement. The data therefore reveal a common control-then-explore sequence, implying that teaching should combine rich, pattern-based initial prompts with guidance on intentionally alternating prompting styles throughout a project's development.

4.1 RQ1: Common ChatGPT prompt structures used in vibe coding

Table 1: Prompt structures used at the beginning of tasks

VPS1. When I start a vibe-coding task, I usually begin with:					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Structured Prompt	94	78.3	78.3	78.3
	Unstructured Prompt	26	21.7	21.7	100.0
	Total	120	100.0	100.0	

Table 2: Prompt structures used at the beginning of tasks

VPS8. In the last month, my prompting style was mostly:					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Structured Prompt	41	34.2	34.2	34.2
	Unstructured Prompt	40	33.3	33.3	67.5
	Mixed Prompt	39	32.5	32.5	100.0
	Total	120	100.0	100.0	

4.2 RQ2: Influence of AI chatbot lessons/exposure on prompting style

Table 3: One-way ANOVA test and post hoc Tukey test on months of AI chatbot usage

ANOVA						
		Sum of Squares	df	Mean Square	F	Sig.
Prompt Engineering Competency	Between Groups	10.350	3	3.450	8.088	.000
	Within Groups	49.479	116	.427		
	Total	59.828	119			
Vibe-Coding Prompting Style & Usage	Between Groups	6.453	3	2.151	4.695	.004
	Within Groups	53.140	116	.458		
	Total	59.593	119			
Perceived Output Efficiency: Structured vs Unstructured	Between Groups	4.887	3	1.629	6.551	.000
	Within Groups	28.841	116	.249		
	Total	33.728	119			

Post Hoc Tests

Multiple Comparisons							
Tukey HSD							
Dependent Variable	(I) Months using AI chatbots regularly	(J) Months using AI chatbots regularly	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Prompt Engineering Competency	0-3 months	4-6 months	.02543	.16014	.999	-.3920	.4429
		7-9 months	-.61296*	.16177	.001	-1.0346	-.1913
		More than 10 months	-.54058*	.16944	.010	-.9822	-.0989
	4-6 months	0-3 months	-.02543	.16014	.999	-.4429	.3920
		7-9 months	-.63839*	.17304	.002	-1.0894	-.1873
		More than 10 months	-.56601*	.18022	.011	-1.0358	-.0962
	7-9 months	0-3 months	.61296*	.16177	.001	.1913	1.0346
		4-6 months	.63839*	.17304	.002	.1873	1.0894
		More than 10 months	.07238	.18168	.978	-.4012	.5459
	More than 10 months	0-3 months	.54058*	.16944	.010	.0989	.9822
		4-6 months	.56601*	.18022	.011	.0962	1.0358
		7-9 months	-.07238	.18168	.978	-.5459	.4012
Vibe-Coding Prompting Style & Usage	0-3 months	4-6 months	-.01273	.16596	1.000	-.4453	.4199
		7-9 months	-.41736	.16765	.067	-.8544	.0196
		More than 10 months	-.52641*	.17560	.017	-.9841	-.0687
	4-6 months	0-3 months	.01273	.16596	1.000	-.4199	.4453
		7-9 months	-.40463	.17933	.114	-.8721	.0628
		More than 10 months	-.51368*	.18677	.034	-1.0005	-.0268
	7-9 months	0-3 months	.41736	.16765	.067	-.0196	.8544
		4-6 months	.40463	.17933	.114	-.0628	.8721
		More than 10 months	-.10905	.18828	.938	-.5998	.3817

	More than 10 months	0-3 months	.52641*	.17560	.017	.0687	.9841
		4-6 months	.51368*	.18677	.034	.0268	1.0005
		7-9 months	.10905	.18828	.938	-.3817	.5998
Perceived Output Efficiency: Structured vs Unstructured	0-3 months	4-6 months	.06755	.12226	.946	-.2512	.3863
		7-9 months	-.39304*	.12351	.010	-.7150	-.0711
		More than 10 months	-.35256*	.12936	.037	-.6898	-.0154
	4-6 months	0-3 months	-.06755	.12226	.946	-.3863	.2512
		7-9 months	-.46059*	.13211	.004	-.8050	-.1162
		More than 10 months	-.42011*	.13760	.015	-.7788	-.0614
	7-9 months	0-3 months	.39304*	.12351	.010	.0711	.7150
		4-6 months	.46059*	.13211	.004	.1162	.8050
		More than 10 months	.04048	.13871	.991	-.3211	.4020
	More than 10 months	0-3 months	.35256*	.12936	.037	.0154	.6898
		4-6 months	.42011*	.13760	.015	.0614	.7788
		7-9 months	-.04048	.13871	.991	-.4020	.3211

The mean difference is significant at the 0.05 level

Table 4: One-way ANOVA test and post hoc Tukey test on prior prompt engineering exposure

ANOVA						
		Sum of Squares	df	Mean Square	F	Sig.
Prompt Engineering Competency	Between Groups	10.814	3	3.605	8.531	.000
	Within Groups	49.014	116	.423		
	Total	59.828	119			
Vibe-Coding Prompting Style & Usage	Between Groups	.978	3	.326	.645	.588
	Within Groups	58.615	116	.505		
	Total	59.593	119			
Perceived Output Efficiency: Structured vs Unstructured	Between Groups	3.529	3	1.176	4.518	.005
	Within Groups	30.199	116	.260		
	Total	33.728	119			

Post Hoc Tests

Multiple Comparisons							
Tukey HSD							
Dependent Variable	(I) Prior 'prompt engineering' training	(J) Prior 'prompt engineering' training	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Prompt Engineering Competency	Never	Brief (less than 2 hours)	-.36821	.15732	.095	-.7783	.0419
		Workshop (3-6 hours)	-.62575*	.16648	.002	-1.0597	-.1918
		Course Module (more than 6 hours)	-.85504*	.18371	.000	-1.3339	-.3762
	Brief (less than 2 hours)	Never	.36821	.15732	.095	-.0419	.7783
		Workshop (3-6 hours)	-.25754	.15876	.370	-.6714	.1563
		Course Module (more than 6 hours)	-.48683*	.17675	.034	-.9476	-.0261
	Workshop (3-6 hours)	Never	.62575*	.16648	.002	.1918	1.0597
		Brief (less than 2 hours)	.25754	.15876	.370	-.1563	.6714
		Course Module (more than 6 hours)	-.22929	.18495	.603	-.7114	.2528
	Course Module (more than 6 hours)	Never	.85504*	.18371	.000	.3762	1.3339
		Brief (less than 2 hours)	.48683*	.17675	.034	.0261	.9476
		Workshop (3-6 hours)	.22929	.18495	.603	-.2528	.7114
Vibe-Coding Prompting Style & Usage	Never	Brief (less than 2 hours)	-.08855	.17204	.955	-.5370	.3599
		Workshop (3-6 hours)	-.11888	.18205	.914	-.5934	.3557
		Course Module (more than 6 hours)	-.27688	.20090	.516	-.8006	.2468
	Brief (less than 2 hours)	Never	.08855	.17204	.955	-.3599	.5370
		Workshop (3-6 hours)	-.03033	.17361	.998	-.4829	.4222
		Course Module (more than 6 hours)	-.18833	.19329	.764	-.6922	.3155
		Never	.11888	.18205	.914	-.3557	.5934

	Workshop (3-6 hours)	Brief (less than 2 hours)	.03033	.17361	.998	-.4222	.4829
		Course Module (more than 6 hours)	-.15800	.20225	.863	-.6852	.3692
	Course Module (more than 6 hours)	Never	.27688	.20090	.516	-.2468	.8006
		Brief (less than 2 hours)	.18833	.19329	.764	-.3155	.6922
		Workshop (3-6 hours)	.15800	.20225	.863	-.3692	.6852
Perceived Output Efficiency: Structured vs Unstructured	Never	Brief (less than 2 hours)	-.23404	.12349	.236	-.5559	.0878
		Workshop (3-6 hours)	-.37527*	.13068	.025	-.7159	-.0346
		Course Module (more than 6 hours)	-.48003*	.14421	.006	-.8559	-.1041
	Brief (less than 2 hours)	Never	.23404	.12349	.236	-.0878	.5559
		Workshop (3-6 hours)	-.14123	.12462	.670	-.4661	.1836
		Course Module (more than 6 hours)	-.24599	.13874	.292	-.6076	.1157
	Workshop (3-6 hours)	Never	.37527*	.13068	.025	.0346	.7159
		Brief (less than 2 hours)	.14123	.12462	.670	-.1836	.4661
		Course Module (more than 6 hours)	-.10476	.14517	.888	-.4832	.2737
	Course Module (more than 6 hours)	Never	.48003*	.14421	.006	.1041	.8559
		Brief (less than 2 hours)	.24599	.13874	.292	-.1157	.6076
		Workshop (3-6 hours)	.10476	.14517	.888	-.2737	.4832

The mean difference is significant at the 0.05 level

Table 3 presents the effects of months of regular AI chatbot use through one-way ANOVA analyses, while Table 4 summarises prior prompt-engineering training. Months of regular AI chatbot use significantly influenced Prompt Engineering Competency, $F(3,116) = 8.088$, $p < .001$, $\eta^2 = .173$; Vibe-Coding Prompting Style and Usage, $F(3,116) = 4.695$, $p = .004$, $\eta^2 = .108$; and Perceived Output Efficiency, $F(3,116) = 6.551$, $p < .001$, $\eta^2 = .145$. Tukey comparisons showed that students with 7–9 months and more than 10 months of experience scored significantly higher than those with 0–6 months on selected outcomes ($p < .05$).

The 7–9 month and >10-month groups, however, were not statistically distinguishable in this sample. This pattern should be interpreted cautiously, as a

non-significant difference alone does not establish a formal plateau. In addition, weekly AI-tool usage frequency did not show statistically significant effects for Prompt Engineering Competency, $F(3,116) = 0.297$, $p = .827$, $\eta^2 = .008$; Vibe-Coding Prompting Style and Usage, $F(3,116) = 0.379$, $p = .769$, $\eta^2 = .010$; or Perceived Output Efficiency, $F(3,116) = 0.067$, $p = .978$, $\eta^2 = .002$.

Regarding prior training, students who had attended workshops or completed course modules reported higher Prompt Engineering Competency, $F(3,116) = 8.531$, $p < .001$, $\eta^2 = .181$, than students with no training, and Perceived Output Efficiency also differed by training level, $F(3,116) = 4.518$, $p = .005$, $\eta^2 = .105$. However, differences in Vibe-Coding Prompting Style and Usage were not statistically significant, $F(3,116) = 0.645$, $p = .588$, $\eta^2 = .016$. Read together, these findings suggest that sustained exposure and formal instruction are associated with higher self-perceived mastery and efficiency, whereas simple weekly frequency of use is not.

Regarding prior training, students who had participated in workshops or completed course modules demonstrated higher competency than those with no or very brief training, $F(3,116) = 8.531$, $p < .001$ (with Tukey $p \leq .034$). Perceived output efficiency was also greater among trained students, $F(3,116) = 4.518$, $p = .005$ (with selected Tukey $p \leq .025$), although differences in prompting style and usage were not statistically significant. On balance, these findings indicate that sustained experience combined with formal instruction, rather than simply frequent weekly use is linked to higher perceived mastery and effectiveness. This suggests that targeted, structured teaching plays a crucial role in transforming exposure into lasting skills.

4.3 RQ3: Impact of educational backgrounds on prompting style

Table 5: One-way ANOVA test on the students' previous education backgrounds

ANOVA						
		Sum of Squares	df	Mean Square	F	Sig.
Prompt Engineering Competency	Between Groups	1.963	2	.981	1.984	.142
	Within Groups	57.866	117	.495		
	Total	59.828	119			
Vibe-Coding Prompting Style & Usage	Between Groups	3.464	2	1.732	3.611	.030
	Within Groups	56.129	117	.480		
	Total	59.593	119			
Perceived Output Efficiency: Structured vs Unstructured	Between Groups	.960	2	.480	1.714	.185
	Within Groups	32.768	117	.280		
	Total	33.728	119			

Table 6: Post Hoc Tukey HSD Test in comparing previous education backgrounds with the factors

Multiple Comparisons							
Tukey HSD							
Dependent Variable	(I) Previous education background	(J) Previous education background	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Prompt Engineering Competency	Humanities/Arts	Technical/STEM	-.20490	.16224	.419	-.5900	.1802
		Social Sciences	-.29689	.15171	.128	-.6570	.0633
	Technical/STEM	Humanities/Arts	.20490	.16224	.419	-.1802	.5900
		Social Sciences	-.09199	.16058	.835	-.4732	.2892
	Social Sciences	Humanities/Arts	.29689	.15171	.128	-.0633	.6570
		Technical/STEM	.09199	.16058	.835	-.2892	.4732
Vibe-Coding Prompting Style & Usage	Humanities/Arts	Technical/STEM	-.38854*	.15979	.043	-.7679	-.0092
		Social Sciences	-.32392	.14942	.081	-.6786	.0308
	Technical/STEM	Humanities/Arts	.38854*	.15979	.043	.0092	.7679
		Social Sciences	.06463	.15815	.912	-.3108	.4401
	Social Sciences	Humanities/Arts	.32392	.14942	.081	-.0308	.6786
		Technical/STEM	-.06463	.15815	.912	-.4401	.3108
Perceived Output Efficiency: Structured vs Unstructured	Humanities/Arts	Technical/STEM	-.14538	.12209	.461	-.4352	.1444
		Social Sciences	-.20714	.11416	.169	-.4782	.0639
	Technical/STEM	Humanities/Arts	.14538	.12209	.461	-.1444	.4352
		Social Sciences	-.06176	.12084	.866	-.3486	.2251
	Social Sciences	Humanities/Arts	.20714	.11416	.169	-.0639	.4782
		Technical/STEM	.06176	.12084	.866	-.2251	.3486

The mean difference is significant at the 0.05 level

Table 5 displays the results of one-way ANOVAs comparing students' previous educational backgrounds (Humanities and Arts, Technical and STEM, and Social Sciences), while Table 6 provides Tukey HSD post hoc comparisons. No significant differences emerged for Prompt Engineering Competency, $F(2,117) = 1.984$, $p = .142$, $\eta^2 = .033$, or Perceived Output Efficiency, $F(2,117) = 1.714$, $p = .185$, $\eta^2 = .028$.

However, Vibe-Coding Prompting Style and Usage differed significantly by educational background, $F(2,117) = 3.611$, $p = .030$, $\eta^2 = .058$. Tukey tests indicated that students from Technical/STEM backgrounds reported significantly higher style-and-usage scores than students from Humanities/Arts backgrounds (mean difference = 0.389, $p = .043$), while other pairwise differences were not significant. Group means were Humanities/Arts = 3.46, Social Sciences = 3.79, and Technical/STEM = 3.85 for Vibe-Coding Prompting Style and Usage.

These findings suggest that educational background influences how students approach vibe-coding prompting styles, though it does not significantly affect their self-perceived capability or beliefs about output efficiency. This pattern points to a common foundation of competency across disciplines, with variations primarily in prompt adoption preferences. Therefore, incorporating discipline-specific examples into teaching materials may encourage wider uptake without the need to create separate skill development tracks for each field.

4.4 RQ4: Influence of gender on prompting style

Table 7: Group means and variation of gender demography

Group Statistics					
	Gender	N	Mean	Std. Deviation	Std. Error Mean
AI Literacy	Male	59	3.5017	.64894	.08448
	Female	61	3.4630	.74749	.09571
Prompt Engineering Competency	Male	59	3.6507	.73199	.09530
	Female	61	3.6203	.69191	.08859
Vibe-Coding Prompting Style & Usage	Male	59	3.6859	.69837	.09092
	Female	61	3.6992	.72227	.09248
Perceived Output Efficiency: Structured vs Unstructured	Male	59	3.9831	.54144	.07049
	Female	61	4.0361	.52663	.06743

Table 8: Independent sample t-test on gender

Independent Samples Test										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
AI Literacy	Equal variances assumed	1.251	.266	.303	118	.763	.03874	.12796	-.21466	.29214
	Equal variances not assumed			.303	116.661	.762	.03874	.12766	-.21409	.29158
Prompt Engineering Competency	Equal variances assumed	.852	.358	.233	118	.816	.03035	.12999	-.22707	.28777
	Equal variances not assumed			.233	117.056	.816	.03035	.13011	-.22733	.28803
Vibe-Coding Prompting Style & Usage	Equal variances assumed	.002	.969	-.102	118	.919	-.01325	.12976	-.27021	.24371
	Equal variances not assumed			-.102	118.000	.919	-.01325	.12969	-.27006	.24356
Perceived Output Efficiency: Structured vs Unstructured	Equal variances assumed	.166	.685	-.544	118	.588	-.05301	.09750	-.24609	.14006
	Equal variances not assumed			-.543	117.558	.588	-.05301	.09755	-.24619	.14016

Table 7 displays group means and variation, while Table 8 shows the results of independent-samples t-tests comparing male and female students. No statistically significant gender differences were found for AI Literacy, $t(118) = 0.303$, $p = .763$, $d = 0.06$; Prompt Engineering Competency, $t(118) = 0.233$, $p = .816$, $d = 0.04$; Vibe-Coding Prompting Style and Usage, $t(118) = -0.102$, $p = .919$, $d = 0.02$; or Perceived Output Efficiency, $t(118) = -0.544$, $p = .588$, $d = 0.10$. Group means were closely aligned across male and female students, indicating minimal practical as well as statistical differences.

5. Discussion

5.1 RQ1: Common ChatGPT prompt structures used in vibe-coding

Students demonstrate a clear pattern: they usually begin with structured prompts and then shift to using a nearly equal mix of structured, unstructured, and mixed styles as tasks progress. This reflects a "control-then-explore" cycle, where roles, constraints, and examples establish a stable foundation for initial outputs, followed by iterative experimentation. Existing research sheds light on this behaviour: a pattern-based approach provides reusable prompt structures that enhance reliability (White et al., 2023), while novices relying on random edits often face inconsistencies (Zamfirescu-Pereira et al., 2023). Studies on AI-supported writing further reveal that students use structure to manage tone and coherence while negotiating their authorial voice (Wang, 2024). Our findings align with this literature: structured prompts reduce fragility, and blending styles supports ongoing refinement. Teaching students to start with pattern-rich prompts and then deliberately iterate is not contradictory but rather a productive sequence; pattern literacy establishes a secure baseline from which exploration can enrich nuance without losing clarity.

5.2 RQ2: Influence of AI-chatbot lessons and exposure on prompting style

Greater experience, particularly around seven to nine months or more, and participation in formal lessons such as workshops or modules were associated with higher self-perceived competency and perceived efficiency, while weekly frequency of use did not show statistically significant effects. This pattern is consistent with evidence that explicit instruction and prior learning are more consequential for generative AI literacy than exposure alone (O'Dea et al., 2024; Lee & Palmer, 2025). It also aligns with work emphasising the importance of disciplined prompting and structured guidance in educational uses of large language models (Kasneci et al., 2023; Walter, 2024). Rather than demonstrating that instruction causes mastery, the present cross-sectional findings suggest that scaffolded teaching and sustained experience are linked to stronger self-perceived capability. For practice, this implies that higher-education programmes should prioritise structured prompt-engineering instruction, feedback, and guided experimentation over simple tool access.

5.3 RQ3: Impact of educational background on prompting style

Although educational background did not significantly affect competency or perceived efficiency, it was related to students' reported vibe-coding style and usage, with those from STEM fields tending to adopt these practices more than

peers in Humanities or Arts. Reviews note that research on chatbots in higher education remains fragmented and often lacks common conceptual frameworks (McGrath et al., 2024), which may encourage discipline-specific habits rather than genuine skill disparities. Faculty affiliation can also influence early technical understanding and adoption behaviours (Toker Gokce et al., 2024), which helps explain the stylistic edge seen in STEM students.

Studies of teacher-AI collaboration also suggest that how prompts are used depends more on pedagogical orchestration than discipline alone (Jeon & Lee, 2023). Therefore, this finding reflects differences in adoption preferences rather than unequal ability. The implication is to maintain a core of pattern-based prompting instruction, while tailoring examples to fit disciplinary genres, since shared methods build equivalent competence, and contextualised tasks improve engagement by reflecting differing norms and rhetorical expectations.

5.4 RQ4: Influence of gender on prompting style

No significant gender differences were observed in AI literacy, prompt engineering competency, style and usage, or perceived output efficiency; group means were closely aligned. This mirrors adoption studies where demographic factors lose explanatory power once perceived usefulness and mentoring are considered (Patterson et al., 2024). These results temper concerns about gender gaps in technical subskills by suggesting that differences may fade when participants have equal training and similar task demands (Toker Gokce et al., 2024). These null results are plausible in an environment where students share comparable educational opportunities and challenges. For curriculum development, the implication is clear: focus on delivering inclusive, high-quality instruction rather than segregating by gender because providing equitable scaffolds and chances for practice appears sufficient to foster comparable prompting mastery across diverse groups.

5.5 Limitations

This study has several limitations. First, it relies on self-reported perceptions, which may not fully reflect students' actual prompt-engineering performance. Second, the cross-sectional design does not support causal inference and captures perceptions at only one point in time. Third, the sample was drawn from a single institution and disciplinary context, which limits broader generalizability. Fourth, the use of questionnaire-based composite scores means that the findings are potentially affected by common method bias. Finally, the study reports the available reliability evidence transparently. However, subscale-specific reliability coefficients, formal test–retest coefficients, and full assumption diagnostics were not retained in the available output and therefore could not be reported. Future work should address these gaps through multi-institutional designs, behavioral log data, archived prompts, and performance-based assessments.

5.6 Recommendations for Future Research

Future studies should adopt longitudinal, experimental, or quasi-experimental designs to test how training intensity influences changes in prompting behaviour over time. Researchers should also extend the evidence base by sampling across multiple institutions and disciplines and by triangulating self-reports with

behavioural log data, archived prompts, and rubric-based output evaluation. Comparing pattern-based workshops with more conventional teaching formats may also help identify cost-effective pedagogical strategies for developing prompt-engineering competence in educational app design. Collectively, such approaches would help clarify how AI literacy, prompting behaviour, and perceived efficiency interact in authentic learning environments.

6. Conclusion

The findings reinforce the view that prompting can be taught as a pattern-based practice. Students often begin with structured prompts and later adopt more flexible, iterative approaches. This pattern is consistent with scholarship arguing that prompt patterns can scaffold reliability while still allowing exploration (White et al., 2023; Lee & Palmer, 2025). At the same time, the present study focuses on self-perceived mastery rather than objectively assessed prompt quality. Within that limitation, the results suggest that structured instruction and sustained exposure are associated with stronger confidence and perceived efficiency, whereas gender differences appear negligible and educational background appears to shape style adoption more than overall competency. These findings support the integration of explicit prompt-engineering instruction, guided practice, and discipline-sensitive examples in language and communication curricula.

7. Funding

This work was supported by INTI International University in collaboration with Akademi Pengajian Bahasa, Universiti Teknologi MARA Shah Alam.

8. Acknowledgments

The authors thank the faculty members and students who participated in this study. The authors also thank the institution for the access to resources and technical support that made the research possible.

9. Declarations

Informed consent was obtained from all participants prior to questionnaire completion. Participation was voluntary, responses were anonymised for analysis, and no personal identifiers were collected. This study involved a minimal-risk educational survey of university students and did not involve patients or clinical procedures. In accordance with institutional practice for minimal-risk educational research, formal ethics committee approval was not required. The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

10. Data Availability and Sharing Policy

Data Availability Statement. The anonymised dataset supporting the findings of this study is available from the authors through the following repository link: https://drive.google.com/drive/folders/1efHwjdwGSGhy95hig-XUc_MBqHK-AREH?usp=drive_link. Access is provided in accordance with institutional and ethical requirements for anonymised participant data.

11. References

Ahmad, N., Alias, F. A., Hamat, M., & Mohamed, S. A. (2024). Reliability analysis: Application of Cronbach's Alpha in research instruments. *Journal of Computer and Mathematical Sciences*. https://appspenang.uitm.edu.my/sigcs/2024-2/Articles/20244_ReliabilityAnalysis-ApplicationOfCronbachsAlphaInResearchInstruments.pdf

Borromeo, A. S., Manaloto, A. M., Santos, M. J. M. D., Antonio, R. P., Soyosa, M. D., & Wider, W. (2025). Harnessing generative AI in nursing education: A bibliometric review. *Teaching and Learning in Nursing*. https://doi.org/10.1016/j.teln.2025.04.014

Bujang, M. A., Omar, E. D., Foo, D. H. P., & Hon, Y. K. (2024). Sample size determination for conducting a pilot study to assess reliability of a questionnaire. Restorative Dentistry & Endodontics, 49(1), e3. https://doi.org/10.5395/rde.2024.49.e3

Cain, W. (2024). Prompting change: Exploring prompt engineering in large language model AI and its potential to transform education. *TechTrends, 68*(1), 47–57. https://doi.org/10.1007/s11528-023-00896-0

Fritsch, T. (2024). Chatbots: An overview of current issues and challenges. In *Advances in Information and Communication* (pp. 84–104). Springer. https://doi.org/10.1007/978-3-031-53960-2_7

Giray, L. (2023). Prompt engineering with ChatGPT: A guide for academic writers. *Annals of Biomedical Engineering, 51*(12), 2629–2633. https://doi.org/10.1007/s10439-023-03272-4

Jeon, J., & Lee, S. (2023). Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. *Education and Information Technologies, 28*, 15873–15892. https://doi.org/10.1007/s10639-023-11834-1

Knoth, N., Tolzin, A., Janson, A., & Leimeister, J. M. (2024). AI literacy and its implications for prompt engineering strategies. Computers and Education: Artificial Intelligence, 6, 100225. https://doi.org/10.1016/j.caeai.2024.100225

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Kowang, T. O., Yew, L. K., Fei, G. C., & Hee, O. C. (2025). Determinants of artificial intelligence acceptance among undergraduates. *International Journal of Evaluation and Research in Education*, 14(4). https://doi.org/10.11591/ijere.v14i4.32565

Lee, D., & Palmer, E. (2025). Prompt engineering in higher education: A systematic review to help inform curricula. *International Journal of Educational Technology in Higher Education, 22*, 7. https://educationaltechnologyjournal.springeropen.com/articles/10.1186/s41239-025-00503-7

McGrath, C., Farazouli, A., & Cerratto-Pargman, T. (2024). Generative AI chatbots in higher education: A review of an emerging research area. *Higher Education, 89*, 1533–1549. https://doi.org/10.1007/s10734-024-01288-w

Nyimbili, F., & Nyimbili, L. (2024). Types of purposive sampling techniques with their examples and application in qualitative research studies. *British Journal of Multidisciplinary and Advanced Studies, 5*(1), 90–99. https://doi.org/10.37745/bjmas.2022.0419

O'Dea, X., Ng, D. T. K., O'Dea, M., & Shkuratskyy, V. (2024). Factors affecting university students' generative AI literacy: Evidence and evaluation in the UK and Hong Kong contexts. *Policy Futures in Education*. https://doi.org/10.1177/14782103241287401

OpenAI. (2025). *Best practices for prompt engineering with the OpenAI API*. OpenAI Help Center. https://help.openai.com/en/articles/6654000-best-practices-for-prompt-engineering-with-the-openai-api

Patterson, A., Frydenberg, M., & Basma, L. (2024). Examining generative artificial intelligence adoption in academia: A UTAUT perspective. *Issues in Information Systems, 25*(3), 238–251. https://iacis.org/iis/2024/3iis2024_238-251.pdf

Quamar, A. H., Schmeler, M. R., McCue, M., et al. (2023). Test–retest reliability of the Electronic Instrumental Activities of Daily Living Satisfaction Assessment (EISA): A cohort study. *American Journal of Occupational Therapy, 77*(6), 7706205140. https://doi.org/10.5014/ajot.2023.050285

Rokeman, N. R. M. (2024). Likert measurement scale in education and social sciences: Explored and explained. *EDUCATUM Journal of Social Sciences, 10*(1), 77–89. https://doi.org/10.37134/ejoss.vol10.1.7.2024

Siripipatthanakul, S., Muthmainnah, B., Asrifan, A., et al. (2023). *Quantitative research in education*. ResearchGate. https://www.researchgate.net/publication/369013292QuantitativeResearchinEducation

Toker Gokce, A., Deveci Topal, A., Kolburan Geçer, A., & Eren, C. D. (2024). Investigating the level of artificial intelligence literacy of university students using decision trees. *Education and Information Technologies, 30*, 6765–6784. https://doi.org/10.1007/s10639-024-13081-4

Wang, C. (2024). Exploring students' generative AI-assisted writing processes: Perceptions and experiences from native and nonnative English speakers. *Technology, Knowledge and Learning, 30*, 1825–1846. https://doi.org/10.1007/s10758-024-09744-3

Walter, Y. (2024). Embracing the future of Artificial Intelligence in the classroom: The relevance of AI literacy, prompt engineering, and critical thinking in modern education. International Journal of Educational Technology in Higher Education, 21, 15. https://doi.org/10.1186/s41239-024-00448-3

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. https://doi.org/10.48550/arXiv.2302.11382

Wong, T. A., Tan, K. T. L., Darmaraj, S. R., Loo, J. T. K., & Ng, A. H. H. (2025). Social capital development in online education and its impact on academic performance and satisfaction. *Higher Education, Skills and Work-Based Learning*. https://doi.org/10.1108/heswbl-12-2023-0332

Woo, D. J., Wang, D., Yung, T., & Guo, K. (2024). Effects of a prompt engineering intervention on undergraduate students' AI self-efficacy, AI knowledge, and prompt engineering ability: A mixed methods study. *Education and Information Technologies*. https://arxiv.org/pdf/2408.07302

Yang, Y., Wen, X., & Maidin, S. S. (2024). Generative AI tools in higher education emerging research: A bibliometric analysis of co-citation and co-word analysis. In *Proceedings of the 2024 3rd International Conference on Artificial Intelligence and Education (ICAIE 2024)*. ACM. https://doi.org/10.1145/3722237.3722266

Yusoff, M. S. B., Arifin, W. N., & Hadie, S. N. H. (2021). ABC of questionnaire development and validation for survey research. Education in Medicine Journal, 13(1), 97–108. https://doi.org/10.21315/eimj2021.13.1.10

Zakariya, Y. F. (2022). Cronbach's alpha in mathematics education research: Its appropriateness, overuse, and alternatives. *Frontiers in Psychology, 13*, 1074430. https://doi.org/10.3389/fpsyg.2022.1074430

Zamfirescu-Pereira, J. D., Wong, R. Y., Hartmann, B., & Yang, Q. (2023). Why Johnny can't prompt: How non-AI experts try (and fail) to design LLM prompts. *Proceedings of*

the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23). https://doi.org/10.1145/3544548.3581388

Zhang, D., Yang, H., He, Y., & Guo, W. (2025). Modeling the relationships between secondary school students' AI learning attitude, AI literacy and AI career interest. *Education and Information Technologies*. https://doi.org/10.1007/s10639-025-13715-1

Zhao, D. (2024). The impact of AI-enhanced natural language processing tools on writing proficiency: An analysis of language precision, content summarization, and creative writing facilitation. *Education and Information Technologies, 30*(1), 8055–8086. https://doi.org/10.1007/s10639-024-13145-5