Equivalent: Jurnal Ilmiah Sosial Teknik
Volume 8, Issue 6, 2486-2500
e_ISSN: 2775-0833
https://jurnalequivalent.id/index.php/jequi/index
DOI: doi.org/10.59261/jequi.v8i6.492

Generative AI Continuance Intention Among Indonesian Higher Education Students: Integrating the Information Systems Success Model and the Expectation-Confirmation Model

Kenny William Phan*
Universitas Bina Nusantara, Indonesia

Viany Utami Tjhin
Universitas Bina Nusantara, Indonesia

***Corresponding author:**
Kenny William Phan, Universitas Bina Nusantara, Indonesia.
✉kenny.phan@binus.ac.id

Article Info:
***Article history*:**
Received: Oktober 06, 2026
Revised: November 12, 2026
Accepted: November 14, 2026
Available Online: November 23, 2026

***Keywords*:**
Generative AI;
Continuance Intention;
IS Success Model;

This is an open access article under the CC BY-SA license.

Copyright © 2026 by Author. Published by Politeknik Siber Cerdika International

Abstract

Background: Generative artificial intelligence (GenAI) adoption among Indonesian university students has reached 95%, shifting research attention toward continuance intention. However, existing studies have largely focused on non-Indonesian contexts or specific paid platforms.
Objective: This study integrates the Information Systems Success Model (ISSM) and Expectation-Confirmation Model (ECM) to explain GenAI continuance intention among Indonesian higher education students across multiple platforms.
Methods: An online survey collected data from students with at least three months of academic GenAI experience. After eligibility and response-quality screening, 219 valid responses were analyzed using partial least squares structural equation modeling (PLS-SEM) with 5,000 bootstrap resamples.
Results: Seven of eight hypotheses were supported. Information quality (β = 0.389) and service quality (β = 0.355) significantly influenced confirmation, whereas system quality did not (β = 0.118, p = 0.279). Confirmation influenced perceived usefulness (β = 0.634) and satisfaction (β = 0.217), while perceived usefulness influenced satisfaction (β = 0.642). Satisfaction (β = 0.431) and perceived usefulness (β = 0.265) explained 43.5% of continuance intention variance. Serial mediation confirmed significant indirect effects of information and service quality through the ECM pathway, but not system quality.
Conclusion: Information and service quality are key drivers of GenAI continuance intention, while system quality may represent a baseline expectation. However, construct overlap warrants cautious interpretation.

To cite this article: Phan, K. W., & Tjhin, V. U. (2026). Generative AI continuance intention among Indonesian higher education students: Integrating the Information Systems Success Model and the Expectation-Confirmation Model. *Equivalent: Jurnal Ilmiah Sosial Teknik*, *8*(6), 2486-2500. https://doi.org/10.59261/jequi.v8i6.492

INTRODUCTION

Generative artificial intelligence (GenAI) has rapidly reshaped higher education. Unlike earlier AI applications designed primarily for classification or prediction, GenAI tools built on large language models (LLMs), such as ChatGPT, Gemini, Claude, and Copilot, generate new text, code, and other content in response to user prompts (Kasneci et al., 2023). Students use these tools for concept clarification, summarization, brainstorming, academic writing, and translation (Bahroun et al., 2023). The Chegg Global Student Survey 2025 reported that 80% of university students across 15 countries used GenAI for academic support, with Indonesia recording the highest adoption rate at 95% (Chegg, 2025; GoodStats, 2025a).

Indonesia was also the fifth-largest source of ChatGPT traffic worldwide in August 2025,

accounting for approximately 216 million visits GoodStats (2025b), while a national survey of 1,501 students found that 86.21% used AI at least once a month (Tirto, 2024). The reported 95% figure is treated as a survey-specific indicator rather than a population estimate for all Indonesian university students. It reflects the sampling frame, timing, and operational definition of study-related GenAI use employed by the cited survey. Likewise, ChatGPT web traffic statistics are used only to indicate the service's popularity and not as evidence of the number of unique student users or the intensity of academic use.

These figures indicate that GenAI has evolved from an experimental tool into part of students' everyday learning infrastructure. When adoption approaches saturation, the question of whether students will adopt GenAI becomes less informative than the question of whether, and why, they will continue using it. This post-adoption behavior is captured by the concept of continuance intention Bhattacherjee (2001), and research on GenAI in education has similarly shifted from tool-centered descriptions toward learner-centered concerns (Wu et al., 2025). Nevertheless, Indonesian studies have predominantly examined adoption intention using the Technology Acceptance Model (TAM), the Unified Theory of Acceptance and Use of Technology 2 (UTAUT2), or the Theory of Planned Behavior (Habibi et al., 2023, 2024; Sinaga et al., 2024; Wahdah et al., 2025). These pre-adoption frameworks are less suited to explaining how accumulated usage experiences shape decisions to continue using a system.

Two information systems (IS) frameworks are particularly relevant to post-adoption behavior. The IS Success Model (ISSM) proposes that system quality, information quality, and service quality influence system use, user satisfaction, and net benefits DeLone & McLean (1992) and has been validated across a wide range of IS contexts (Jeyaraj, 2020). The Expectation-Confirmation Model (ECM), derived from expectation-confirmation theory Oliver (1980), explains continuance intention as an outcome of confirmation, post-adoption perceived usefulness, and satisfaction Bhattacherjee (2001), with its explanatory power supported by meta-analytic evidence (Ambalov, 2018). The two frameworks are complementary: ISSM specifies which system attributes matter but does not explain the psychological processes through which users evaluate them, whereas ECM explains those processes but does not specify the system attributes that drive confirmation. Integrating ISSM and ECM therefore offers a more comprehensive explanation of post-adoption behavior. Such integrated models have been proposed in the context of virtual classrooms Alshammari & Alshammari (2024), while an integration of the Technology Acceptance Model and ISSM has been applied to GenAI continuance intention among Chinese students (Wang, 2025).

In the context of GenAI, the three quality dimensions have specific implications. System quality refers to response speed, reliability, interface clarity, and adaptability across tasks. Information quality refers to the accuracy, relevance, completeness, and currency of generated outputs, which are particularly important because LLMs can produce plausible but inaccurate content (Kasneci et al., 2023). Service quality refers to response consistency, the handling of follow-up questions, availability, and the overall smoothness of interaction, given that GenAI services are delivered largely through conversational interfaces (DeLone & McLean, 2003). Because students repeatedly experience these attributes through their interactions with GenAI tools, they represent potential antecedents of expectation confirmation.

Empirical research on GenAI continuance intention is growing. In a study involving 282 GenAI users, Jung & Jo (2025) found that confirmation strongly influenced perceived usefulness and satisfaction, which, in turn, shaped continuance intention. Kavitha & Joshith (2025) identified satisfaction as the most critical determinant of continuance intention toward ChatGPT among 350 Indian Generation Z students, while Bekturova et al. (2025) confirmed the roles of perceived usefulness and satisfaction among 334 Kazakhstani students using ChatGPT for academic writing. Studies grounded in ISSM, however, have reported inconsistent findings regarding the effects of the three quality dimensions.

Berényi et al. (2025) found that information quality and system quality, but not service quality, significantly influenced the outcomes examined among 466 Hungarian and Romanian students, whereas Aziz et al. (2025) found significant effects for all three dimensions among 1,019 Pakistani students. Mun & Hwang (2025) further demonstrated that information quality and information usefulness drive continued ChatGPT use through source trust. In Indonesia, Nabila et

al. (2026) reported that all three quality dimensions influenced perceived usefulness and price fairness, which, in turn, strengthened continuance intention among 200 ChatGPT Premium subscribers.

Three gaps emerge from this literature. First, there is a contextual-theoretical gap: GenAI continuance intention has been studied primarily in China, India, Eastern Europe, Pakistan, Kazakhstan, and Korea, whereas Indonesia, despite its highest reported adoption rate, has received relatively little research attention. Second, there is an object-of-study gap: the only Indonesian ISSM-based study focused on paid subscribers of a single tool, leaving the broader population of students who use GenAI tools across different platforms, including free versions, largely unexamined. Third, there is a theoretical gap: no Indonesian study has examined the full serial mediation chain through which system quality, information quality, and service quality influence continuance intention via confirmation, perceived usefulness, and satisfaction. The novelty claim is based on a targeted review of recent studies on GenAI continuance intention identified through combinations of the search terms 'generative AI', 'continuance intention', 'expectation confirmation', 'IS success', and 'higher education'. The review focused on peer-reviewed studies available to the authors and was used to map their geographical settings, user groups, platform specificity, and theoretical frameworks. Because this was not a preregistered systematic review, the claim is framed as a literature-based research gap rather than an exhaustive assertion that no other Indonesian studies exist. Beyond geographical coverage, prior evidence has not yet established whether the three ISSM quality dimensions retain distinct explanatory roles when incorporated into the ECM chain in a heterogeneous, multi-platform GenAI setting. This issue is particularly important because quality, confirmation, perceived usefulness, and satisfaction may exhibit empirical overlap.

This study addresses these gaps and makes three contributions. First, it examines the broader population of Indonesian higher education students who use GenAI as a general category rather than focusing on a specific tool or subscription tier. Second, it integrates ISSM and ECM into a single serial mediation model and decomposes the specific indirect effects through which each quality dimension influences continuance intention. Third, it contextualizes user satisfaction in relation to academic outcomes, namely, students' perceived improvements in grades and understanding of course material. Accordingly, this study aims to analyze the effects of system quality, information quality, and service quality on confirmation; the effects of confirmation on perceived usefulness and satisfaction; the effects of perceived usefulness on satisfaction and continuance intention; and the effect of satisfaction on continuance intention. It also aims to validate the integrated ISSM–ECM model in the Indonesian higher education context. A testable proposition is that information quality and service quality are indirectly associated with continuance intention through confirmation, perceived usefulness, and satisfaction, whereas the role of system quality is treated as an empirical question rather than assumed to be negligible in a mature market.

The research model (Figure 1) positions the three ISSM quality dimensions as exogenous antecedents of confirmation, the entry point of the ECM mechanism. When students perceive GenAI as responsive, reliable, and easy to interact with (system quality), as producing accurate, relevant, complete, and current outputs (information quality), and as handling follow-up questions consistently and smoothly (service quality), their actual experience is more likely to meet or exceed prior expectations (Nabila et al., 2026). Hence, H1: system quality positively influences confirmation; H2: information quality positively influences confirmation; and H3: service quality positively influences confirmation.

Following ECM, confirmed expectations lead users to revise their usefulness beliefs upward and generate a positive affective response Oliver (1980), relationships also observed for GenAI (Wang, 2025). Usefulness further increases satisfaction Alshammari & Babu (2025), and both usefulness and satisfaction drive continuance intention (Nabila et al., 2026). Hence, H4: confirmation positively influences perceived usefulness; H5: confirmation positively influences satisfaction; H6: perceived usefulness positively influences satisfaction; H7: perceived usefulness positively influences continuance intention; and H8: satisfaction positively influences continuance intention.

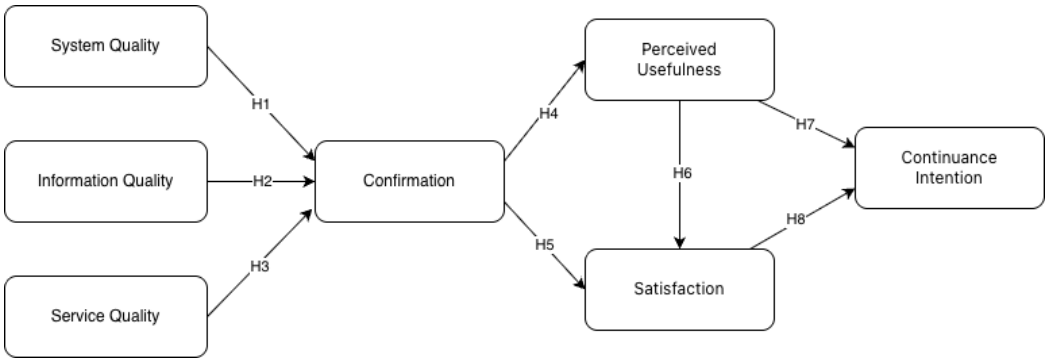

Figure 1. Research model
Source: Adapted from (DeLone & McLean, 2003)

METHOD

This study employed a quantitative, explanatory, cross-sectional survey design. Partial least squares structural equation modeling (PLS-SEM) was selected because it is suitable for prediction-oriented models involving multiple mediating constructs and has been widely applied in recent studies of GenAI continuance intention (Hair Jr et al., 2021). The seven constructs were measured using items adapted from validated scales: system quality, information quality, and service quality from DeLone & McLean (2003), contextualized with reference to Berényi et al. (2025); confirmation from Jung & Jo (2025); perceived usefulness from Davis (1989); satisfaction from Oliver (1980); and continuance intention from (Wang, 2025). The satisfaction construct was extended with two items measuring GenAI's perceived impact on academic grades (SAT4) and understanding of course material (SAT5) to connect affective evaluations with academic outcomes. All items were measured using a six-point Likert scale ranging from 1 (strongly disagree) to 6 (strongly agree).

This format eliminates the neutral midpoint and may reduce central-tendency bias (Abdul, 2010). The questionnaire consisted of three sections: an informed consent statement and three screening questions reflecting the inclusion criteria; nine demographic and usage-related items (including tools used, primary tool, frequency, and duration of use); and the construct measurement items. The questionnaire was administered in both Indonesian and English and pilot-tested with 40 eligible respondents. The items retained for analysis in the final model are presented in Table 2. A six-point scale was selected to elicit directional responses and maintain consistency across the adapted measurement instrument. Although eliminating the neutral midpoint may reduce habitual midpoint selection, it may also compel genuinely neutral respondents to select a non-neutral response. This trade-off is therefore acknowledged as a measurement consideration rather than treated solely as an advantage.

The target population comprised active students enrolled at public (*perguruan tinggi negeri*, PTN) and private (*perguruan tinggi swasta*, PTS) higher education institutions in Indonesia, including diploma, undergraduate, master's, and doctoral students. Purposive sampling was employed based on three inclusion criteria: current enrollment at an Indonesian higher education institution, use of at least one GenAI tool for academic purposes, and a minimum of three months of GenAI use. The minimum usage period was intended to ensure sufficient experience for comparing initial expectations with actual usage (Bhattacherjee, 2001). An a priori minimum sample size calculation was conducted using Cohen (1992) multiple regression power framework ($\alpha = 0.05$, power = 0.80, and assumed $R^2 = 0.25$), based on the largest number of predictors for any endogenous construct in the model (three).

The calculation indicated a minimum required sample size of 59 respondents. This threshold addresses statistical power only and is not considered evidence of measurement model validity. GenAI was operationalized as a category of generative tools used for academic purposes rather than as a single homogeneous platform. Because ChatGPT, Gemini, Claude, Copilot, Perplexity AI, DeepSeek, and other tools differ in their interfaces, model versions, and access tiers, the primary platform used by each respondent was recorded descriptively, and platform heterogeneity was treated as a boundary condition for interpreting the findings.

Data were collected through Google Forms over a four-week period using recruitment

channels that included Instagram, X, LinkedIn, WhatsApp groups, and university student communities. Respondents could optionally complete a separate form to participate in a small giveaway, thereby preserving the anonymity of the main survey. Of the 280 responses received, 22 did not meet the inclusion criteria. A straight-lining check based on the standard deviation of each respondent's responses across all Likert-scale items resulted in the exclusion of a further 39 responses with SD ≤ 0.30, including 11 responses in which identical answers had been provided to every item. The final analytical sample therefore comprised 219 respondents, representing a retention rate of 78.2%.

Data were analyzed using SmartPLS 4 Ringle et al. (2022) in two stages, following the approach outlined by (Hair Jr et al., 2021). The measurement model was assessed using outer loadings (≥ 0.70), Cronbach's alpha and composite reliability (ρA and ρC ≥ 0.70), average variance extracted (AVE ≥ 0.50) Fornell & Larcker (1981), and the heterotrait–monotrait ratio (HTMT ≤ 0.85, or ≤ 0.90 for conceptually similar constructs (Henseler et al., 2015). The structural model was assessed using the variance inflation factor (VIF < 5), path coefficients evaluated through bootstrapping with 5,000 subsamples and two-tailed significance testing, the coefficient of determination (R^2, with values of 0.25, 0.50, and 0.75 indicating weak, moderate, and substantial explanatory power, respectively), effect sizes (f^2, with values of 0.02, 0.15, and 0.35 indicating small, medium, and large effects, respectively; Cohen (1992), and the standardized root mean square residual (SRMR). Specific and total indirect effects were estimated to examine the proposed serial mediation relationships.

All participants were provided with an informed consent statement explaining the study's purpose, the voluntary nature of participation, anonymity, the academic use of the collected data, and their right to withdraw before submitting their responses. Consent was required before participants could proceed to the screening questions. No personally identifiable information was collected through the main survey, and the data were stored in encrypted cloud storage accessible only to the researchers.

RESULTS AND DISCUSSION

Results

Respondent Profile

Table 1 summarizes the respondent profile. The sample was predominantly female (66.2%), aged 22–25 years (55.7%), and enrolled in undergraduate programs (79.9%), with a relatively balanced distribution between public (53.0%) and private (47.0%) institutions. ChatGPT was the most reported primary tool (52.1%), followed by Claude (17.4%) and Gemini (16.9%). The use of multiple tools was common: 74.0% of respondents had used ChatGPT, 58.0% Gemini, 37.0% Claude, 33.8% Perplexity AI, 22.8% Copilot, and 17.4% DeepSeek. Most respondents used GenAI three to five times per week (44.3%), and 44.7% had used GenAI for more than one year. The primary GenAI tool item was single-choice and included an "Other" category, whereas the separate tools-used item allowed multiple selections. Therefore, the two variables serve different descriptive purposes and should not be interpreted as interchangeable measures.

Table 1
Respondent profile (n = 219)

Characteristic	Category	n	%
Gender	Female / Male	145 / 74	66.2 / 33.8
Age	<18 / 18–21 / 22–25 / 26–30 / >30 years	1 / 58 / 122 / 26 / 12	0.5 / 26.5 / 55.7 / 11.9 / 5.5
Education level	Diploma / Undergraduate / Master's / Doctoral	19 / 175 / 24 / 1	8.7 / 79.9 / 11.0 / 0.5
Institution type	Public (PTN) / Private (PTS)	116 / 103	53.0 / 47.0
Field of study	Business, economics & management / STEM / Social sciences & humanities / Health & medicine / Arts & design / Others	86 / 53 / 35 / 25 / 9 / 11	39.3 / 24.2 / 16.0 / 11.4 / 4.1 / 5.0
Primary GenAI tool	ChatGPT / Claude / Gemini / Copilot / DeepSeek / Perplexity AI / Others	114 / 38 / 37 / 14 / 8 / 7 / 1	52.1 / 17.4 / 16.9 / 6.4 / 3.7 / 3.2 / 0.5
Use per week	1–2 / 3–5 / 6–10 times / Daily / Multiple times daily	16 / 97 / 58 / 35 / 13	7.3 / 44.3 / 26.5 / 16.0 / 5.9
Duration of use	3–6 months / 7–12 months / >1 year	59 / 62 / 98	26.9 / 28.3 / 44.7

Source: Survey data (2026)

Measurement Model

Table 2 presents the item means and outer loadings. All items had relatively high mean scores (4.87–5.33 on a six-point scale). Twenty-two of the 28 indicators had outer loadings above 0.70, whereas six indicators (IQ4, PU4, SAT5, SVQ3, SYQ4, and SYQ5) had loadings ranging from 0.592 to 0.697. Following Hair Jr et al. (2021), these indicators were retained because their loadings were statistically significant ($p < 0.001$), they were conceptually essential to their respective constructs, and their removal did not materially improve the average variance extracted (AVE).

Table 2
Measurement items, means, and outer loadings

Item	Statement (English version)	Mean	Loading
SYQ1	GenAI responds to my queries quickly	5.19	0.761
SYQ2	GenAI is reliable for academic activities	5.12	0.719
SYQ3	GenAI is easy to interact with	5.33	0.759
SYQ4	GenAI adapts well to various academic tasks	5.10	0.644
SYQ5	The GenAI interface is clear and easy to understand	5.22	0.592
IQ1	GenAI provides accurate information for my academic needs	5.01	0.792
IQ2	GenAI information is relevant to my coursework	5.19	0.728
IQ3	GenAI provides complete answers to my questions	5.05	0.700
IQ4	Information from GenAI is up to date	4.87	0.664
SVQ1	GenAI provides consistent responses across my queries	5.15	0.705
SVQ2	GenAI handles clarification or follow-up questions effectively	5.08	0.739
SVQ3	GenAI is available whenever I need it	5.28	0.609
SVQ4	GenAI delivers a smooth user experience	5.01	0.720
CON1	My experience with GenAI is better than I initially expected	5.02	0.788
CON2	The benefits I receive from GenAI exceed my initial expectations	4.95	0.786
CON3	Most of my expectations about GenAI have been confirmed	5.11	0.792
PU1	Using GenAI improves my academic productivity	5.27	0.712
PU2	GenAI helps me complete academic tasks more efficiently	5.29	0.718
PU3	GenAI enhances the quality of my academic work	5.27	0.708
PU4	Overall, GenAI is useful for my studies	5.07	0.665

SAT1	I am satisfied with my overall experience using GenAI	5.24	0.781
SAT2	I am pleased with the help GenAI provides for my studies	5.14	0.702
SAT3	My decision to use GenAI was a wise one	5.28	0.815
SAT4	GenAI use has positively impacted my academic grades	5.06	0.713
SAT5	GenAI use has improved my understanding of course materials	5.10	0.697
CI1	I intend to continue using GenAI for my academic activities	5.22	0.825
CI2	I will continue using GenAI regularly in the future	5.01	0.798
CI3	I plan to recommend GenAI to my peers for academic use	4.96	0.809

Note: SYQ = system quality; IQ = information quality; SVQ = service quality; CON = confirmation; PU = perceived usefulness; SAT = satisfaction; CI = continuance intention. All loadings significant at p < 0.001.

As shown in Table 3, composite reliability (ρC) exceeded 0.70 for all constructs (0.788–0.860). Cronbach's α ranged from 0.641 to 0.796, with CON, IQ, PU, and SVQ slightly below 0.70, and AVE ranged from 0.483 to 0.657, with PU, SVQ, and SYQ marginally below 0.50. A sensitivity analysis using the sample before straight-lining removal (n = 258) produced α ≥ 0.72 and AVE ≥ 0.55 for all constructs and the same structural conclusions (seven of eight hypotheses supported). Because Cronbach's α assumes equal indicator loadings and therefore provides a conservative lower-bound estimate, ρA and ρC are the preferred reliability criteria in PLS-SEM Hair Jr et al. (2021), and ρC exceeded 0.70 for every construct, although ρA remained below 0.70 for CON, PU, and SVQ. The marginal values in the final sample are partly attributable to the removal of straight-lining respondents, which reduced artificial response uniformity.

Table 3
Construct reliability and convergent validity

Construct	Cronbach's α	ρA	ρC	AVE
System quality (SYQ)	0.737	0.752	0.825	0.488
Information quality (IQ)	0.694	0.702	0.813	0.522
Service quality (SVQ)	0.641	0.642	0.788	0.483
Confirmation (CON)	0.697	0.698	0.832	0.622
Perceived usefulness (PU)	0.654	0.655	0.794	0.491
Satisfaction (SAT)	0.796	0.800	0.860	0.552
Continuance intention (CI)	0.739	0.741	0.852	0.657

Source: SmartPLS 4 output (n = 219)

Table 4 presents the heterotrait–monotrait ratio (HTMT) results. Twelve of the 21 construct pairs had HTMT values below 0.85, and four had values between 0.85 and 0.90. Five pairs exceeded 0.90: SVQ–CON (0.981), SYQ–SVQ (0.975), IQ–CON (0.961), PU–CON (0.935), and SAT–PU (1.080). These values indicate substantial empirical overlap among the three quality dimensions and confirmation, as well as between satisfaction and perceived usefulness. The constructs were nevertheless retained as theoretically distinct because they are grounded in established frameworks DeLone & McLean (2003), all had composite reliability values above 0.70, and they played differentiated roles in the structural model, as illustrated by the contrasting effects of information and service quality versus system quality on confirmation. The implications of this overlap are discussed further in the Discussion section. Notably, the SAT–PU HTMT value of 1.080 indicates that satisfaction and perceived usefulness are not empirically distinguishable in this sample. This overlap may partly reflect the extended satisfaction items concerning academic grades and course understanding. Accordingly, strong claims regarding discriminant validity and the distinctiveness of the PU–SAT mechanism are not warranted, and the serial mediation results should be treated as provisional pending item refinement and model re-estimation.

Table 4
Heterotrait–monotrait ratio (HTMT)

	SYQ	IQ	SVQ	CON	PU	SAT
IQ	0.896					
SVQ	0.975	0.878				
CON	0.835	0.961	0.981			
PU	0.722	0.786	0.853	0.935		
SAT	0.693	0.668	0.765	0.836	1.080	
CI	0.669	0.683	0.710	0.731	0.864	0.829

Source: SmartPLS 4 output (n = 219)

Structural Model

Inner variance inflation factor (VIF) values ranged from 1.000 to 2.546, indicating no critical collinearity among the predictor constructs. Table 5 and Figure 2 show that seven of the eight hypotheses were supported. Information quality (β = 0.389, p < 0.001) and service quality (β = 0.355, p = 0.001) significantly influenced confirmation, whereas system quality did not (β = 0.118, p = 0.279; f^2 = 0.015). Confirmation had a strong effect on perceived usefulness (β = 0.634, f^2 = 0.674) and a smaller effect on satisfaction (β = 0.217, f^2 = 0.077). Perceived usefulness also had a strong effect on satisfaction (β = 0.642, f^2 = 0.675). Both satisfaction (β = 0.431, f^2 = 0.129) and perceived usefulness (β = 0.265, f^2 = 0.049) significantly influenced continuance intention.

Table 5
Hypothesis testing results

H	Path	β	SE	t	p	f^2	Decision
H1	SYQ → CON	0.118	0.109	1.082	0.279	0.015	Not supported
H2	IQ → CON	0.389	0.086	4.542	<0.001	0.189	Supported
H3	SVQ → CON	0.355	0.110	3.213	0.001	0.149	Supported
H4	CON → PU	0.634	0.052	12.109	<0.001	0.674	Supported
H5	CON → SAT	0.217	0.068	3.202	0.001	0.077	Supported
H6	PU → SAT	0.642	0.059	10.830	<0.001	0.675	Supported
H7	PU → CI	0.265	0.100	2.656	0.008	0.049	Supported
H8	SAT → CI	0.431	0.099	4.352	<0.001	0.129	Supported

Note: SE = bootstrap standard error (β/t); bootstrapping with 5,000 subsamples, two-tailed.

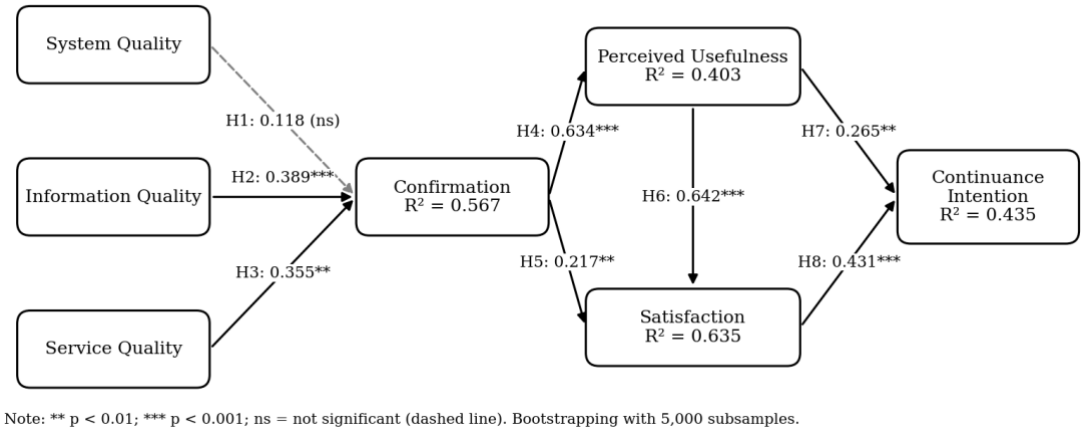

Figure 2. Structural model results
Source: SmartPLS 4 output (n = 219), redrawn by the authors

The model explained 56.7% of the variance in confirmation (adjusted R^2 = 0.561), 40.3% in perceived usefulness (adjusted R^2 = 0.400), 63.5% in satisfaction (adjusted R^2 = 0.632), and 43.5% in continuance intention (adjusted R^2 = 0.430). These results indicate moderate explanatory power for confirmation and satisfaction and weak-to-moderate explanatory power

for perceived usefulness and continuance intention (Hair Jr et al., 2021). The standardized root mean square residual (SRMR) was 0.076 for the saturated model and 0.091 for the estimated model, while the normed fit index (NFI) was 0.641.

Mediation Analysis

Table 6 reports selected specific and total indirect effects. The strongest specific indirect effect was observed for CON → PU → SAT (β = 0.407, p < 0.001), which exceeded the direct CON → SAT effect (β = 0.217), indicating complementary partial mediation. The complete serial mediation pathway, CON → PU → SAT → CI, was significant (β = 0.176), as was PU → SAT → CI (β = 0.277). Information quality and service quality were associated with continuance intention through the complete serial pathway (β = 0.068 and 0.062, respectively), whereas all six indirect pathways originating from system quality were non-significant. The largest total indirect effect on continuance intention originated from confirmation (β = 0.438), followed by information quality (β = 0.170) and service quality (β = 0.155).

Table 6

Selected specific and total indirect effects

Path	β	t	p	Result
Specific indirect effects				
CON → PU → SAT	0.407	8.204	<0.001	Significant
PU → SAT → CI	0.277	4.556	<0.001	Significant
CON → PU → SAT → CI	0.176	4.285	<0.001	Significant
CON → PU → CI	0.168	2.546	0.011	Significant
CON → SAT → CI	0.094	2.093	0.036	Significant
IQ → CON → PU	0.247	4.678	<0.001	Significant
IQ → CON → PU → SAT → CI	0.068	3.109	0.002	Significant
IQ → CON → SAT → CI	0.036	1.797	0.072	Not significant
SVQ → CON → PU	0.225	2.861	0.004	Significant
SVQ → CON → PU → SAT → CI	0.062	2.323	0.020	Significant
SVQ → CON → PU → CI	0.060	1.961	0.050	Borderline
SYQ → CON → PU	0.075	1.069	0.285	Not significant
SYQ → CON → PU → SAT → CI	0.021	1.022	0.307	Not significant
Total indirect effects on CI				
CON → CI	0.438	9.150	<0.001	Significant
PU → CI	0.277	4.556	<0.001	Significant
IQ → CI	0.170	4.219	<0.001	Significant
SVQ → CI	0.155	2.846	0.004	Significant
SYQ → CI	0.052	1.035	0.301	Not significant

Source: SmartPLS 4 output (n = 219), bootstrapping with 5,000 subsamples

Discussion

Table 7 compares the key findings with those of the studies that informed the hypotheses. The core relationships proposed by the expectation-confirmation model (ECM) were consistently supported across contexts, whereas the significance of the information systems success model (ISSM) quality dimensions varied across populations. Because the outcomes examined for the quality dimensions differed across studies (confirmation in this study; use and satisfaction in Berényi et al. (2025); satisfaction, attitude, and performance in Aziz et al. (2025) and perceived usefulness and price fairness in Nabila et al. (2026) the comparison focuses on patterns of statistical significance rather than coefficient magnitudes.

Table 7
Comparison of key findings with prior GenAI studies

Path	This study (Indonesia, n = 219)	(Berényi et al., 2025) (Hungary–Romania, n = 466)	(Aziz et al., 2025) (Pakistan, n = 1,019)	(Nabila et al., 2026) (Indonesia, Premium, n = 200)	(Jung & Jo, 2025) (n = 282)
SYQ → outcome	NS	S	S	S	–
IQ → outcome	S	S	S	S	–
SVQ → outcome	S	NS	S	S	–
CON → PU	S	–	–	–	S
PU → SAT	S	–	–	–	S
SAT → CI	S	–	–	–	S

Note: S = significant; NS = not significant; – = not tested. Source: compiled by the authors.

System Quality as a Baseline Expectation

The most distinctive finding is the non-significant effect of system quality on confirmation. This result differs from those of Nabila et al. (2026), who found system quality to be significant among Pakistani students and Indonesian ChatGPT Premium subscribers, respectively, but is consistent with Berényi et al. (2025), who found service quality, rather than system quality, to be non-significant among Hungarian and Romanian students. Taken together, these studies suggest that one of the three ISSM quality dimensions may lose its unique explanatory power in GenAI contexts and that the dimension exhibiting this pattern may depend on the user population.

Three explanations are plausible. First, the quality dimensions overlap empirically (HTMT for SYQ–SVQ = 0.975; SYQ–IQ = 0.896). Consequently, after controlling for information and service quality, the unique variance attributable to system quality becomes negligible (f^2 = 0.015), consistent with the high intercorrelations among ISSM quality dimensions reported by (Jeyaraj, 2020). Second, mainstream GenAI tools now offer consistently fast, stable, and easy-to-use interfaces, as reflected in the high mean scores for the system quality items (5.10–5.33). When technical performance is perceived as consistently high across tools, system quality may function as a baseline expectation rather than a differentiating factor. However, this interpretation remains a hypothesis requiring direct empirical testing. Third, unlike the sample examined by Nabila et al. (2026), this study was not restricted to paid subscribers, whose financial investment may raise their expectations of technical performance. Because subscription status was not measured in this study, this explanation should also be examined in future research.

The Primacy of Information and Service Quality

Information quality was the strongest antecedent of confirmation, consistent with the findings of (Mun & Hwang, 2025). Academic use of GenAI is output-oriented: students rely on answers, explanations, and summaries. Moreover, the well-documented risk of inaccurate or fabricated content Dwivedi et al. (2023) makes accuracy and relevance central to whether users' expectations are met. Notably, information currency (IQ4) had the lowest mean score (4.87) and an outer loading below the recommended threshold, suggesting that timeliness is the weakest aspect of perceived information quality. This finding is plausible because many GenAI tools rely on training data with a knowledge cutoff, meaning that students may encounter outdated information when working on current topics.

Service quality was also statistically significant, in contrast to the findings of (Berényi et al., 2025). In this study, service quality encompassed response consistency, handling of follow-up questions, availability, and smoothness of interaction. Its significance suggests that Indonesian students value the conversational, multi-turn nature of GenAI, possibly because academic tasks are often refined iteratively. However, the very high HTMT value between service quality and confirmation (0.981) indicates that part of this association may reflect overlapping perceptions.

Therefore, the relative magnitudes of the H2 and H3 coefficients should be interpreted cautiously. Availability (SVQ3) also had the weakest outer loading within the service quality construct, suggesting that students may regard constant availability as an expected feature and evaluate service quality primarily through the quality of conversational interactions.

The ECM Mechanism and Determinants of Continuance

The core ECM relationships were strongly supported, with large effects of confirmation on perceived usefulness and perceived usefulness on satisfaction, consistent with (Kavitha & Joshith, 2025). The mediation results indicate that perceived usefulness is a central component of the modeled mechanism: confirmation of expectations is associated with greater satisfaction primarily through stronger usefulness perceptions (CON → PU → SAT), although confirmation also has a smaller direct effect on satisfaction. However, despite the statistical significance of the indirect effects, the absence of discriminant validity between perceived usefulness and satisfaction prevents a strong claim that usefulness constitutes a distinct primary psychological mechanism. Therefore, the mediation evidence should be interpreted as indicative of an interrelated post-adoption evaluation process rather than as definitive confirmation of a fully differentiated serial mediation mechanism.

The mediation pattern also clarifies the relationship between ISSM and ECM. Information and service quality influence continuance intention only indirectly, through confirmation and predominantly through perceived usefulness. The total indirect effect of confirmation on continuance intention (β = 0.438) was larger than any other total indirect effect, positioning confirmation as a central component of the integrated model. Moreover, the pathway IQ → CON → SAT → CI was not statistically significant, whereas IQ → CON → PU → SAT → CI was significant. These findings suggest that quality perceptions are associated with satisfaction primarily through stronger usefulness perceptions rather than through a direct association with positive affect. This interpretation is consistent with Bhattacherjee (2001) view of post-acceptance usefulness as the cognitive link between confirmation and satisfaction.

Satisfaction was a stronger predictor of continuance intention than perceived usefulness, consistent with Kavitha & Joshith (2025) and the ECM proposition that satisfaction is the most proximal determinant of continuance intention (Bhattacherjee, 2001). Thus, students may continue using GenAI not only because they perceive it as useful but also because they evaluate their overall experience positively. At the same time, the HTMT value of 1.080 between satisfaction and perceived usefulness indicates that respondents did not empirically distinguish these constructs. This overlap is partly consistent with the ECM framework, in which satisfaction is influenced by perceived usefulness. However, the additional satisfaction items concerning academic grades and understanding (SAT4 and SAT5) may also have brought the satisfaction construct closer to perceived usefulness. Consequently, the relative magnitudes of the H7 and H8 coefficients should be interpreted as indicative rather than conclusive.

Implications and Limitations

Theoretically, this study demonstrates that ISSM and ECM can complement one another in explaining GenAI continuance intention: ISSM identifies the quality attributes associated with confirmation, whereas ECM explains how confirmation is associated with continuance intention through perceived usefulness and satisfaction. The study adds evidence from a broad population of Indonesian higher education students rather than exclusively from paid subscribers of a single tool. It also supports the view that the relative importance of ISSM quality dimensions is context-dependent. In this sample, technical quality may function as a baseline condition. However, because market maturity was not directly measured, the non-significant system quality pathway should be regarded as a context-specific empirical finding requiring replication rather than evidence of a general shift associated with market maturity.

Practically, GenAI developers seeking to retain Indonesian student users should prioritize the accuracy, currency, and consistency of outputs, as well as the effective handling of follow-up interactions, rather than focusing exclusively on further improvements in technical performance, which students already perceive as relatively high. Given the widespread adoption reflected in this sample, higher education institutions should shift their focus from whether to permit GenAI

use to how to support its responsible and sustained use. For example, institutions could integrate students' ability to critically evaluate GenAI outputs into coursework and issue clear guidelines on ethical use. Students would also benefit from developing realistic expectations and critically verifying GenAI-generated outputs, as confirmation and perceived usefulness are central components of the modeled process leading to satisfaction. For policymakers, these findings provide an empirical basis for developing national guidelines on GenAI use in higher education, particularly those emphasizing information quality standards and AI literacy.

This study has several limitations. First, its cross-sectional design captures continuance intention rather than actual long-term usage behavior. Second, purposive online sampling may overrepresent technologically engaged students, and the sample is skewed toward female and business students. Third, the measurement model does not fully satisfy conventional assessment criteria: six outer loadings were below 0.70, three AVE values were marginally below 0.50, several HTMT values exceeded 0.90, and the SRMR for the estimated model (0.091) exceeded the commonly used reference value of 0.08. Finally, GenAI was treated as a general category, and respondents' subscription status was not measured.

CONCLUSION

This study integrated the information systems success model (ISSM) and the expectation-confirmation model (ECM) to examine GenAI continuance intention among 219 Indonesian higher education students. Seven of the eight structural hypotheses were supported. Information quality and service quality were significantly associated with confirmation, whereas system quality was not. Confirmation was strongly associated with perceived usefulness, and both perceived usefulness and satisfaction were significantly associated with continuance intention. These findings suggest that, in this sample, the perceived quality of GenAI outputs and service interactions was more closely associated with the modeled continuance process than system-level quality alone. System quality may function as a baseline expectation, although this explanation requires further empirical investigation.

The serial mediation estimates were statistically significant for the information- and service-quality pathways. However, the discriminant validity problems, particularly the substantial empirical overlap between perceived usefulness and satisfaction, prevent definitive conclusions about the distinctiveness of the proposed psychological stages. Accordingly, the findings support the potential value of an integrated ISSM–ECM perspective while highlighting the need for further measurement refinement in GenAI contexts. Future studies should revise overlapping measurement items, test alternative construct specifications, employ longitudinal research designs, and compare student populations across institutions and countries before drawing stronger conclusions about the proposed serial mediation mechanisms. The conclusions are limited to this cross-sectional, purposively sampled group of 219 students and concern continuance intention rather than actual long-term usage behavior. Longitudinal research, instrument revalidation, and multigroup comparisons by primary platform, field of study, institution type, and subscription status are needed before stronger causal or generalizable claims can be established.

ACKNOWLEDGEMENT

The authors would like to express their sincere gratitude to Universitas Bina Nusantara, Indonesia, for its academic support and encouragement throughout this research. Appreciation is also extended to all individuals who provided valuable insights and constructive feedback during the research and manuscript preparation.

AUTHOR CONTRIBUTION STATEMENT

Kenny William Phan: Conceptualization, Methodology, Investigation, Data curation, Formal analysis, Visualization, Writing – original draft. Viany Utami Tjhin: Supervision, Conceptualization, Methodology (including the extension of the satisfaction construct), Validation, Writing – review and editing. All authors have read and approved the final manuscript.

REFERENCES

Abdul. (2010). Quality of Psychology Test Between Likert Scale 5 and 6 Points. *Journal of Social Sciences*, *6*(3), 399–403. https://doi.org/10.3844/jssp.2010.399.403

Alshammari, S. H., & Alshammari, R. A. (2024). An integration of expectation confirmation model and information systems success model to explore the factors affecting the continuous intention to utilise virtual classrooms. *Scientific Reports*, *14*(1), 18491. https://doi.org/10.1038/s41598-024-69401-8

Alshammari, S. H., & Babu, E. (2025). The mediating role of satisfaction in the relationship between perceived usefulness, perceived ease of use and students' behavioural intention to use ChatGPT. *Scientific Reports*, *15*(1), 7169. https://doi.org/10.1038/s41598-025-91634-4

Ambalov, I. A. (2018). A meta-analysis of IT continuance: An evaluation of the expectation-confirmation model. *Telematics and Informatics*, *35*(6), 1561–1571. https://doi.org/10.1016/j.tele.2018.03.016

Aziz, F., Li, C., & Khan, A. U. (2025). Transforming behavioral intention and academic performance: ChatGPT-4.0 insights through SEM, ANN, and cIPMA analysis. *Information Development*, *41*(3), 933–956. https://doi.org/10.1177/02666669251336560

Bahroun, Z., Anane, C., Ahmed, V., & Zacca, A. (2023). Transforming Education: A Comprehensive Review of Generative Artificial Intelligence in Educational Settings through Bibliometric and Content Analysis. *Sustainability*, *15*(17), 12983. https://doi.org/10.3390/su151712983

Bekturova, M., Tulepova, S., & Zhaitapova, A. (2025). Predicting Kazakhstani TEFL students' continuance intention towards using ChatGPT in academic writing. *Education and Information Technologies*, *30*(9), 11889–11909. https://doi.org/10.1007/s10639-024-13306-6

Berényi, L., Lázár, E., & Madaras, S. (2025). ChatGPT Acceptance Among Students: An Information System View. *Education Sciences*, *15*(11), 1464. https://doi.org/10.3390/educsci15111464

Bhattacherjee, A. (2001). Understanding Information Systems Continuance: An Expectation-Confirmation Model1. *MIS Quarterly*, *25*(3), 351–370. https://doi.org/10.2307/3250921

Chegg. (2025). Global student survey 2025. Chegg, Inc. https://www.chegg.com/about/research/global-student-survey

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 319–340. https://doi.org/10.2307/249008

DeLone, W. H., & McLean, E. R. (1992). Information systems success: The quest for the dependent variable. *Information Systems Research*, *3*(1), 60–95.

DeLone, W. H., & McLean, E. R. (2003). The DeLone and McLean model of information systems success: a ten-year update. *Journal of Management Information Systems*, *19*(4), 9–30. https://doi.org/10.1080/07421222.2003.11045748

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., … Wright, R. (2023). Opinion Paper: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, *71*, 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research*, *18*(1), 39. https://doi.org/10.2307/3151312

GoodStats. (2025a, October 6). 95% mahasiswa RI gunakan AI dalam proses pembelajaran. https://data.goodstats.id/statistic/95-mahasiswa-ri-gunakan-ai-dalam-proses-pembelajaran-FIm7A

GoodStats. (2025b). Indonesia jadi pengguna ChatGPT tertinggi ke-5 di dunia. https://goodstats.id/article/indonesia-jadi-pengguna-chatgpt-tertinggi-ke5-di-dunia-

mjoj8
Habibi, A., Muhaimin, M., Danibao, B. K., Wibowo, Y. G., Wahyuni, S., & Octavia, A. (2023). ChatGPT in higher education learning: Acceptance and use. *Computers and Education: Artificial Intelligence*, *5*, 100190. https://doi.org/10.1016/j.caeai.2023.100190
Habibi, A., Mukminin, A., Octavia, A., Wahyuni, S., Danibao, B. K., & Wibowo, Y. G. (2024). ChatGPT acceptance and use through UTAUT and TPB: A big survey in five Indonesian Universities. *Social Sciences & Humanities Open*, *10*, 101136. https://doi.org/10.1016/j.ssaho.2024.101136
Hair Jr, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2021). *A primer on partial least squares structural equation modeling (PLS-SEM)*. Sage publications.
Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, *43*(1), 115–135. https://doi.org/10.1007/s11747-014- 0403-8
Jeyaraj, A. (2020). DeLone & McLean models of information system success: Critical meta-review and research directions. *International Journal of Information Management*, *54*, 102139. https://doi.org/https://doi.org/10.1016/j.ijinfomgt.2020.102139
Jung, Y. M., & Jo, H. (2025). Understanding Continuance Intention of Generative AI in Education: An ECM-Based Study for Sustainable Learning Engagement. *Sustainability*, *17*(13), 6082. https://doi.org/10.3390/su17136082
Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … 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
Kavitha, K., & Joshith, V. P. (2025). Exploring Gen Z's Satisfaction and Continuance Intention Towards ChatGPT 4o in Higher Education: PLS-SEM Insights Using the ECM. *Journal of Educational Technology Systems*, *54*(2), 407–430. https://doi.org/10.1177/00472395251380294
Mun, I. B., & Hwang, K.-H. (2025). Understanding ChatGPT continuous usage intention: The role of information quality, information usefulness, and source trust. *Information Development*, *41*(3), 675–691. https://doi.org/10.1177/02666669241307595
Nabila, A. J., Pebrianti, W., Giriati, Heriyadi, & Setiawan, H. (2026). Determinants of Continuance Intention Toward ChatGPT Premium Among Indonesian University Students: An Information Systems Success Model Perspective for Higher Education Management. *Journal of Educational Management Research*, *5*(2). https://doi.org/10.61987/jemr.v5i2.1892
Oliver, R. L. (1980). A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions. *Journal of Marketing Research*, *17*(4), 460. https://doi.org/10.2307/3150499
Ringle, C. M., Wende, S., & Becker, J.-M. (2022). SmartPLS 4. SmartPLS GmbH. https://www.smartpls.com
Sinaga, J. N., Panjaitan, E. S., & Nurjanah, S. (2024). Analysis of Factors Affecting the Use of ChatGPT at Mikroskil University: A Study Based on the Extended UTAUT2 Model. *Brilliance: Research of Artificial Intelligence*, *4*(1), 151–161. https://doi.org/10.47709/brilliance.v4i1.3862
Tirto. (2024, May 31). Penggunaan AI di dunia pendidikan makin marak dan merata. https://tirto.id/penggunaan-ai-di-dunia-pendidikan-makin-marak-dan-merata-gZax
Wahdah, S. I., Kurniawan, R. Y., & Irawan, N. (2025). Adoption of ChatGPT in Higher Education: Insights from the Unified Theory of Acceptance and Use of Technology Model. *Indonesian Journal on Learning and Advanced Education (IJOLAE)*, 312–327. https://doi.org/10.23917/ijolae.v7i2.9743
Wang, J. (2025). Students' Continuance Intention to Use Generative AI in Higher Education: Integrating the Technology Acceptance Model and Information Success Model. *Proceedings of the 2025 3rd International Conference on Educational Knowledge and Informatization*, 115–124. https://doi.org/10.1145/3765325.3765346
Wu, F., Dang, Y., & Li, M. (2025). A Systematic Review of Responses, Attitudes, and Utilization

Behaviors on Generative AI for Teaching and Learning in Higher Education. *Behavioral Sciences*, *15*(4), 467. https://doi.org/10.3390/bs15040467