



Working Paper 2026
- Vol. 2, No. 2

TÁC ĐỘNG CỦA NIỀM TIN VÀO NĂNG LỰC BẢN THÂN VÀ SIÊU NHẬN THỨC ĐẾN KẾT QUẢ HỌC TẬP: VAI TRÒ TRUNG GIAN CỦA TRƯỞNG TÁC VỚI TRÍ TUỆ NHÂN TẠO (AI)

**Nguyễn Tấn Đạt¹, Bùi Tuấn Dũng, Vũ Việt Dũng,
Nguyễn Đỗ Khôi Nguyên, Huỳnh Lê Đình Trung**

Sinh viên Khóa 62 – CLC1 Quản trị Kinh doanh Quốc tế
Trường Đại học Ngoại thương Phân hiệu tại TP. Hồ Chí Minh, Việt Nam

ThS Đào Quốc phương

Giảng viên Phân hiệu

Trường Đại học Ngoại thương Phân hiệu tại TP. Hồ Chí Minh, Việt Nam

Tóm tắt:

Trước đây, giáo dục truyền thống lấy con người làm trung tâm trong quá trình tạo sinh tri thức. Người học sử dụng khả năng tư duy của mình phát triển tri thức mới. Tuy nhiên, với sự bùng nổ của Trí tuệ Nhân tạo Tạo sinh (Generative AI), trật tự ngày nay đã có sự thay đổi. Điều này dẫn đến một câu hỏi lớn trong giáo dục đại học: Liệu trí tuệ nhân tạo có thể thay thế tư duy của con người? Và làm thế nào để người học gìn giữ tính chủ thể trong trí tuệ ấy? Đi sâu vào giải quyết câu hỏi này, nghiên cứu phân tích tác động của niềm tin năng lực bản thân (self-efficacy) và siêu nhận thức (metacognition) đến kết quả học tập của sinh viên khi sử dụng AI. Nghiên cứu đã thu thập dữ liệu 302 sinh viên tại TP. Hồ Chí Minh và sử dụng mô hình cấu trúc tuyến tính từng phần (PLS-SEM). Kết quả cho thấy niềm tin năng lực bản thân, siêu nhận thức và tương tác với AI đều có độ tin cậy và độ giá trị cao (CR đạt 0.70 và AVE đạt 0.50). Đáng chú ý, tương quan cao (0.861) giữa tương tác AI và kết quả học tập phản ánh nguy cơ ngộ nhận giữa thành thạo công cụ với việc làm chủ kiến thức. Đây là cơ sở để các nghiên cứu trong tương lai tiếp tục khám phá các giải pháp nhằm bảo vệ nuôi dưỡng năng lực tư duy của người học trong thời đại GenAI.

Từ khoá: Generative AI, academic self-efficacy, higher education, metacognition, perceived learning.

¹ k62.2311215003@ftu.edu.vn

THE IMPACT OF SELF-EFFICACY AND METACOGNITION ON LEARNING OUTCOMES: THE MEDIATING ROLE OF AI INTERACTION

Abstract:

Conventionally, the form of education has placed human intellectuals on the centre stage of knowledge formation. It was accepted that only learners were able to use their minds in order to create or advance knowledge. The reversal of that hierarchy occurred with the introduction of Generative Artificial Intelligence (GenAI). So, the big question is: is it possible to substitute human thinking with AI? And what about defending the critical thinking ability of students when an AI continues to encroach on the role of the intelligent person? These research questions guide our study of the influence of self-efficacy and metacognition on the achievement of students in AI-intensive learning environments. In this research, we gathered data from 302 undergrads in Ho Chi Minh City, and performed PLS -SEM. The findings demonstrate that the constructs of self-efficacy, metacognition, and AI interaction are strong in nature; CRs of 0.70 and AVEs of 0.50 demonstrate that they are trustworthy and valid. It is important to note that AI interaction has a strong relationship with the perceived learning (0.861), thus there is the likelihood that students are baffled by the way AI tactics can be effectively applied. This leaves us with good empirical ground to explore, in further research, how to safeguard and nurture the cognitive capacity of the learners in the GenAI age.

Keywords: Generative AI, academic self-efficacy, higher education, metacognition, perceived learning.

1. Introduction

Hence, greater education has been constructed on the notion that it is only humans who construct knowledge (Bandura, 2001). Learning would imply there is indeed deep thought involved, but with a generative AI, it is more than a tool. It is rather similar to a system that attempts to imitate the way humans learn and develop knowledge. GenAI systems are capable of creating knowledge like humans. That is what makes us ask: is the material AI generates true knowledge or a smooth-caliber simulation? As a matter of fact, the greatest difficulty is not merely connecting technology, but ensuring that we, as the learners, continue to hone our brains. To achieve that, we require internal competencies such as academic self-efficacy and metacognition to improve interrelating with GenAI.

To demonstrate that the hypothesis has a cause and a consequence, we should see that these notions are in fact distinct. It is the reason why this paper is all about establishing a sound measurement foundation. Based on the responses of 302 undergraduates in Ho Chi Minh City, the researcher will: (i) demonstrate that the entire structural system is grounded on sound and valid psychometric measures; (ii) demonstrate that meta-cognitive monitoring is distinct, not identical to a metacognitive knowledge; and (iii) provide future research with an empirical evidence to determine how to minimize that illusion of competence when incorporating AI.

The remainder of this paper is structured as follows: Section 2 establishes the theoretical framework; Section 3 details the measurement validation approach; and Section 4 and 5 present the empirical results and their implications for sustainable AI adoption.

2. Literature Review

2.1. Key concepts and theoretical background

2.1.1. Self-efficacy

Self-efficacy reflects learners' beliefs about their capability to perform academic tasks and manage learning demands. Within self-regulated learning (thereafter called SRL), self-efficacy is a key motivational belief that shapes goal commitment, effort, persistence (Pintrich, 1999; Schunk & Zimmerman, 2008). Empirically, academic self-efficacy is associated with greater use of metacognitive learning strategies and improved performance, including in online settings (Hayat & Shateri, 2019; Yokoyama, 2024).

2.1.2. Metacognition

Metacognition is the awareness and regulation of one's thinking through planning, monitoring, and evaluation (Flavell, 1979; Brown, 1987). A central distinction separates metacognitive knowledge from metacognitive monitoring (Flavell, 1987; Schraw & Moshman, 1995).

Metacognitive knowledge enables intentional strategy selection and justification (Schraw, 1998; Schraw et al., 2006). In AI contexts, metacognitive knowledge plausibly extends to knowledge of AI affordances and limits and conditional knowledge about effective AI use (Xiao et al., 2025; Zhai et al., 2025).

Metacognitive monitoring refers to ongoing assessment of comprehension, confidence, and progress; it is expressed through metacognitive judgments like confidence and judgments of learning (Flavell, 1979; Schraw, 2009). When monitoring weakens in AI-mediated learning, learners are more vulnerable to overreliance and degraded learning processes (Buçinca et al., 2021; Klingbeil et al., 2024; Fan et al., 2024).

Taken together, metacognitive knowledge provides the repertoire and conditional rules for choosing strategies, whereas metacognitive monitoring provides the moment-to-moment calibration that determines whether strategies should be maintained, adapted, or abandoned.

Figure 1 synthesizes the literature review's claim that metacognition (knowledge + monitoring) enables calibrated GenAI use: learners plan prompts/strategies using metacognitive knowledge and verify AI outputs through monitoring to adjust reliance toward task goals (Flavell, 1987; Schraw & Moshman, 1995; Buçinca et al., 2021).

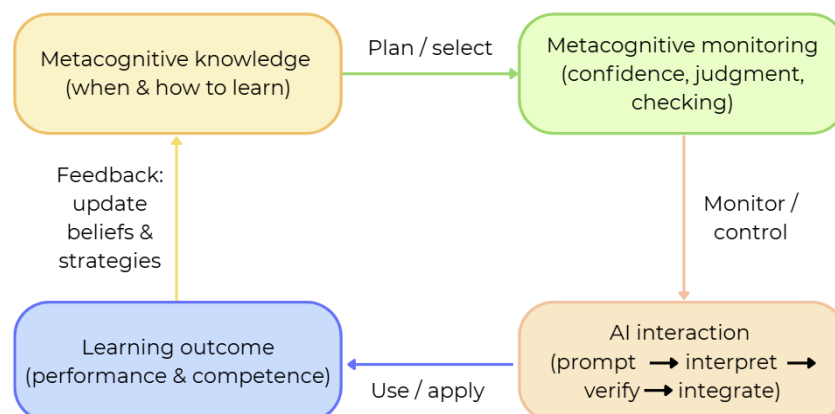


Figure 1. AI-assisted metacognitive calibration cycle (conceptual synthesis).

Source: *The Author (2026)*.

2.1.3. The impact of AI interaction quality on learning outcomes

In AI-assisted contexts, these components jointly shape when learners consult AI, how they formulate prompts, and how rigorously they scrutinize and integrate AI responses. Accordingly, AI interaction quality can be understood as a behavioral manifestation of metacognitive regulation in a hybrid human-AI learning environment (Buçinca et al., 2021; Fan et al., 2024; Zhai et al., 2025).

This study quantifies AI interaction quality as how well learners prompt, interpret, and verifies AI help. This definition emphasizes quality over mere frequency of use. Users may “trust rather than think” when AI makes decisions feel easy (Buçinca et al., 2021), and overreliance can impose performance costs when AI is wrong (Klingbeil et al., 2024). In education, generative AI may induce “metacognitive laziness” if essential thinking is outsourced (Fan et al., 2024), whereas structured AI use that elicits reflection can support learning (Zhai et al., 2025). With AI increasingly embedded in learning, outcomes depend on whether AI usage sustains metacognitive engagement or replaces it via offloading (Buçinca et al., 2021; Fan et al., 2024; Zhai et al., 2025).

Figure 2 illustrates the monitoring–reliance boundary highlighted in the preceding paragraphs: strong monitoring supports selective reliance via verification, whereas weak monitoring paired with high reliance risks adopting fluent but incorrect AI outputs (Klingbeil et al., 2024).

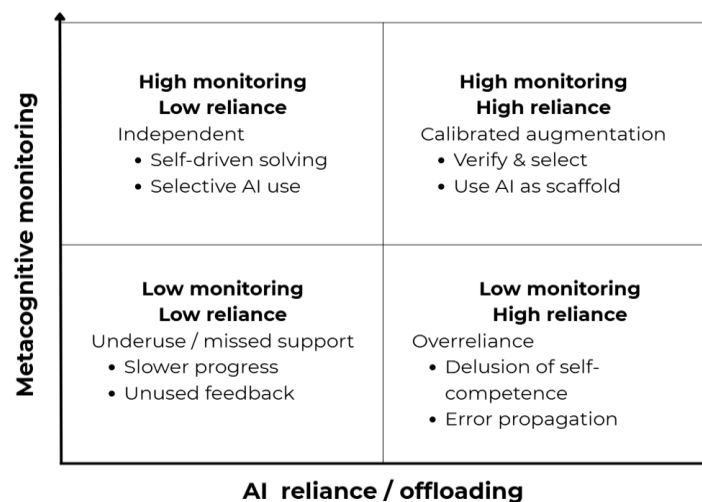


Figure 2. When AI helps vs hurts: monitoring × reliance (conceptual synthesis).

Source: Authors’ synthesis based on Klingbeil et al. (2024).

2.2. Theoretical Framework

2.2.1. Social Cognitive Theory (SCT)

This study is theoretically grounded in Social Cognitive Theory (SCT) (Bandura, 1986, 1997), which conceptualizes learning as the result of reciprocal interactions among personal factors, behaviors, and environmental conditions. Self-efficacy is a key belief shaping effort, persistence, and strategy selection (Bandura, 1997). SCT posits that self-efficacy functions as

a key personal antecedent that activates metacognitive processes, including knowledge of learning strategies and ongoing monitoring of task performance (Schunk & DiBenedetto, 2020).

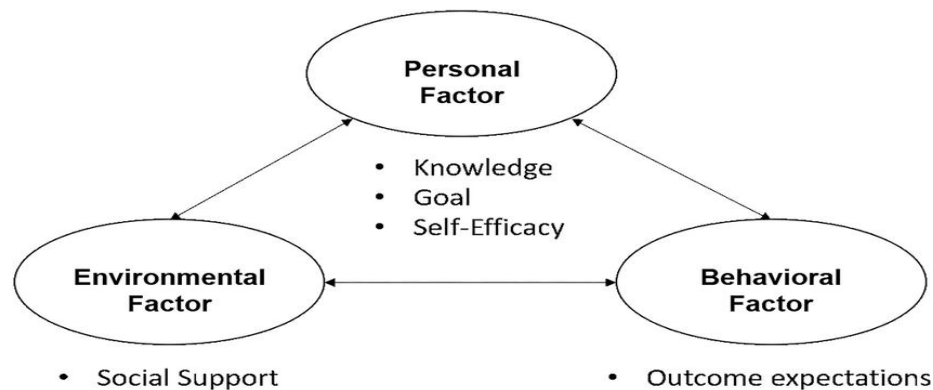


Figure 3. The social cognitive theory and the interaction factors.

Source: Schunk and DiBenedetto (2020).

2.2.2. Zimmerman's Cyclical Model of SRL

This study adopts Zimmerman's (2000) cyclical model of self-regulated learning to explain the procedural mechanism underlying students' interaction with AI. Academic self-efficacy is positioned within the Forethought phase as a key motivational antecedent that initiates regulatory processes (Zimmerman, 2002). In this paper, metacognitive awareness and monitoring represent self-observation, while adaptive AI interaction is conceptualized as a form of environmental self-control. The quality of these performance-phase processes ultimately determines learning outcomes, which constitute the output of the SRL cycle.

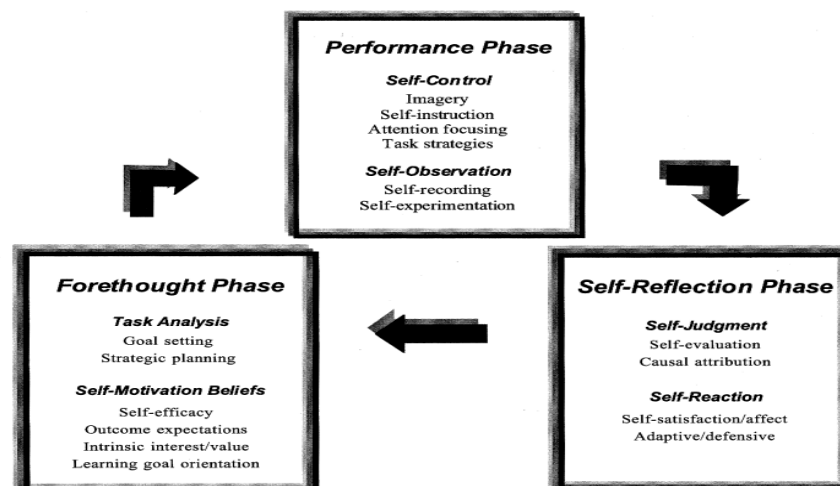


Figure 4. Zimmerman's cyclical phases of self-regulated learning.

Source: Zimmerman (2000).

2.2.3. The structure of Metacognition

Inspired by the influential contributions of Flavell (1979) and Schraw and Dennison (1994), the current research presents the conceptualization of metacognition as a multidimensional variable, that is, having two constituents: Knowledge of Cognition and Regulation of Cognition. Metacognitive awareness, depicted as Knowledge of Cognition, involves declarative and conditional knowledge of learners about strategies and task requirements. Cognition Regulation, which is also referred to as Metacognitive Monitoring,

refers to the dynamic executive processes that help to monitor and observe learning during performance.

2.2.4. Hybrid Human-AI Regulation

In order to capture the complexity of the concept of AI-supported learning, this work uses adaptive AI interactions as part of Hybrid Human-AI Regulation (HHAIR; Molenaar, 2022). The AI provides adaptive scaffolding, such as feedback and reflection cues, and the learners are able to self-regulate using metacognitive reasoning in this hybrid regulatory loop.

3. Research Methodology

3.1. Hypothesis Development and Proposed Research Model

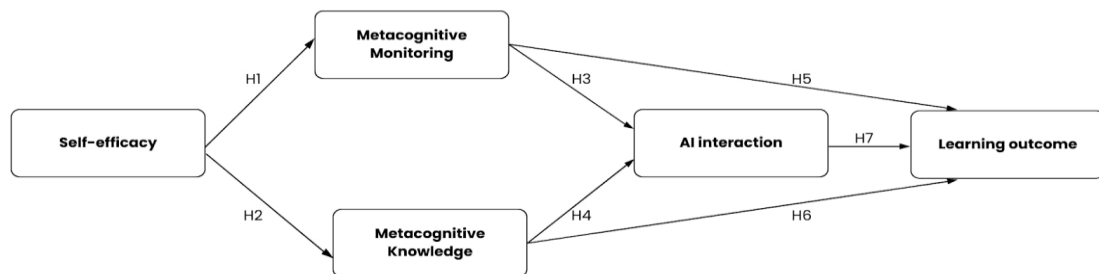


Figure 5. Proposed research model

Sources: The Author (2026)

The proposed model frames learning in the GenAI era as a flow from internal motivation to perceived outcomes. Central to this is Academic Self-Efficacy, the key antecedent driving learner agency. Crucially, AI Interaction is treated not as a technical skill but as a behavioral outcome of metacognition. We posit that adaptive GenAI interaction stems from strong metacognitive monitoring and knowledge. Consequently, the framework proposes a dual pathway to Learning Outcomes: a direct influence from metacognitive engagement and an indirect one mediated by AI interaction. This structure captures the dynamics of hybrid human-AI regulation, where internal regulation guides external algorithmic support.

To operationalize these structural relationships for empirical testing, specific paths were identified based on prior literature. Table 1 synthesizes these study hypotheses and their theoretical rationale, aligned with the proposed research model.

Table 1. Proposed hypotheses and structural paths

Sources: The Author (2026)

Hypothesis	Structural path	Key theoretical backing
H1	Self-efficacy → Metacognitive monitoring	Hayat & Shateri, 2019; Halmo et al., 2024; Schunk & DiBenedetto, 2020
H2	Self-efficacy → Metacognitive knowledge	Schraw et al., 2006; Rahimi & Abedi, 2014
H3	Metacognitive monitoring	Buçinca et al., 2021;

Hypothesis	Structural path	Key theoretical backing
	→ AI interaction	Klingbeil et al., 2024
H4	Metacognitive knowledge → AI interaction	Xiao et al., 2025
H5	Metacognitive monitoring → Learning outcome	Thiede et al., 2003; Dignath et al., 2023; Fan et al., 2024
H6	Metacognitive knowledge → Learning outcome	Coutinho & Neuman, 2008; Teng & Yang, 2023
H7	AI interaction → Learning outcomes	Fernandes et al., 2026; Zhai et al., 2025

3.2. Sampling and Data Collection

In order to experimentally examine the connections between self-efficacy, self-regulated learning, AI interaction, and learning outcomes, this study used a quantitative research approach.

The target population comprised undergraduate students currently enrolled in universities across Ho Chi Minh City, Vietnam. To ensure the relevance of the study to the research context, specific inclusion criteria were established: participants were required to be undergraduates ranging from their first to fourth academic year and must actively utilize Generative AI chatbots- specifically platforms such as ChatGPT, Gemini, or Grok - to support their academic performance.

A minimum sample size of 302 respondents was targeted to ensure sufficient statistical power for data analysis. The data collection process was administered electronically; a structured questionnaire hosted on Google Forms was distributed to prospective participants via institutional email channels, ensuring a streamlined and accessible approach to gathering responses from the target demographic.

3.3. Measurement Scales

3.3.1. Self-Efficacy

To assess the antecedent variable of the research model, Academic Self-Efficacy was measured using a scale adapted from established instruments in educational psychology (Schwarzer, R., & Jerusalem, M., 1995). This concept is defined as students' confidence in their capacity to organize and execute the tasks required to meet predetermined academic goals. The scale consisted of 9 items designed to evaluate the respondents' confidence in mastering complex course materials and their persistence when facing academic challenges.

Participants rated the items on a Likert scale of 5-point (from 1 = Strongly Disagree to 5 = Strongly Agree), where higher scores indicate a greater sense of academic agency.

3.3.2. Metacognition

Metacognitive Knowledge and Metacognitive Monitoring were both measured on a 9-item scale derived on the basis of the highly scaled Metacognitive Awareness Inventory (MAI) by Schraw and Dennison (1994). This framework was chosen as it has strong abilities to measure declarative, procedural, and conditional knowledge of learners in relation to their cognitive processes. Also, it has been found to be reliable in eliciting the necessary regulatory processes, such as planning, information management, and self-monitoring, in the process of learning activities. By means of the use of this classic scale, the active control of cognitive processes by the students can be determined accurately. Through the established theoretical basis, the study will be able to conduct a rigorous assessment of the way students perceive their learning strategies. Everything was determined by having responses on a 5-point Likert scale (from 1 = Strongly Disagree to 5 = Strongly Agree).

3.3.3. AI Interaction

To assess the quality and depth of student engagement with Generative AI, this study utilized a scale adapted from the HAI-capability assessment rubrics (Feiming Li et al, 2025). The original performance-based descriptors were converted into a self-report instrument measured on a 5-point Likert scale (from 1 = Strongly Disagree to 5 = Strongly Agree).

This adaptation rigorously preserved the rubric's core conceptual indicators, ensuring the instrument accurately captures sophisticated prompting strategies, critical evaluation of AI outputs, and the active integration of information, rather than simply measuring the frequency of usage.

3.3.4. Learning Outcomes

In this study, learning outcomes are operationalized as Perceived Learning, defined as the students' self-assessment of their cognitive and affective development. We adopted and modified the Cognitive, Affective, and Psychomotor (thereafter called CAP) Perceived Learning Scale developed by Roandi et al. (2009). The use of a self-report measure is justified by Richmond et al. (1987) and Roandi et al. (2009), suggesting that adult learners can accurately estimate their learning progress, whereas traditional grades may lack consistency across different instructors and fail to capture the multidimensional impact of technological interaction.

While the original CAP scale comprises three domains: Cognitive, Affective, and Psychomotor, this study focuses specifically on the Cognitive and Affective domains to align with the context of AI-assisted learning (excluding the Psychomotor domain, which pertains to physical skills).

The scale consists of 5 items. The wording of original items from Roandi et al. (2009) was slightly modified to specifically refer to the "use of AI tools" during their academic performance. All items are measured using a 5-point Likert scale.

3.4. Analytical Approach

The given research is an empirical test of the proposed research model and assumptions carried out with the help of the Partial Least Squares Structural Equation Modeling (PLS-SEM) based on SmartPLS 4 software. PLS-SEM is selected for its suitability in exploratory research, its robustness in handling non-normal data distributions, and its superior capability in analyzing

complex models involving mediating variables (Hair et al., 2018). The data analysis follows the standard two-stage analytical procedure recommended by Hair et al. (2021).

The first step was focused on evaluating the measurement model in order to test the validity and reliability of the latent constructs. This involved a measure of convergent validity (Average Variance Extracted - AVE), internal consistency reliability, and indicator reliability (outer loadings). Furthermore, discriminant validity was rigorously examined using the Heterotrait-Monotrait (HTMT) ratio, which is considered a superior metric for detecting lack of discriminant validity compared to traditional criteria. While the ultimate goal is to assess structural paths, this working paper prioritizes the rigorous validation of the measurement instruments as a prerequisite for future structural hypothesis testing.

4. Results & Discussion

4.1. Descriptive Statistics

The empirical test is made on 302 valid observations, which are gathered to test the suggested relations between Self-Efficacy (SE), Metacognitive Knowledge (MK), Metacognitive Monitoring (MM), AI Interaction (AI), and Learning Outcomes (LO). The sample size is larger than recommended levels of PLS-SEM, especially when the model has various constructs and indicators, thus making the successive analyses robust and reliable (Hair et al., 2019). The measurement model consists of 45 reflective indicators, which offer sufficient and broad coverage of every latent construct.

Table 2. Descriptive statistics of observed variables

Sources: The Author (2026)

Name	Mean	Median	Standard deviation
SE1	3.877	4.000	0.485
SE2	3.609	4.000	0.944
SE3	3.642	4.000	0.807
SE4	3.695	4.000	0.945
SE5	3.937	4.000	0.965
SE6	3.745	4.000	0.935
SE7	3.831	4.000	0.865
SE8	3.987	4.000	0.898
SE9	3.831	4.000	0.892
MK1	3.957	4.000	0.814
MK2	3.821	4.000	0.918
MK3	3.775	4.000	0.866
MK4	3.616	4.000	0.939

Name	Mean	Median	Standard deviation
MK5	3.993	4.000	0.936
MK6	3.917	4.000	0.996
MK7	3.801	4.000	0.876
MK8	3.781	4.000	0.912
MK9	3.848	4.000	0.932
MM1	3.821	4.000	0.902
MM2	3.884	4.000	0.965
MM3	3.99	4.000	0.957
MM4	3.99	4.000	0.933
MM5	4.01	4.000	0.893
MM6	3.858	4.000	0.915
MM7	3.699	4.000	0.901
MM8	3.891	4.000	0.968
MM9	3.808	4.000	1.038
AI1	4.162	4.000	0.934
AI2	3.96	4.000	0.936
AI3	3.993	4.000	0.824
AI4	4.189	4.000	0.905
AI5	4.142	4.000	0.903
AI6	3.944	4.000	0.834
AI7	4.205	4.000	0.832
AI8	4.01	4.000	0.892
AI9	4.093	4.000	0.804
AI10	3.964	4.000	0.844
AI11	3.93	4.000	0.845
LO1	4.02	4.000	0.874
LO2	3.626	4.000	0.853
LO3	4.026	4.000	0.876
LO4	4.046	4.000	1.018

Name	Mean	Median	Standard deviation
LO5	3.887	4.000	0.789

The descriptive findings present high levels of agreement with all constructs, and the mean scores are primarily above 3.6, and the median value is always 4. This trend indicates positive perceptions of respondents towards the constructs of Self-Efficacy (SE), Metacognitive Knowledge (MK), Metacognitive Monitoring (MM), AI Interaction (AI), as well as Learning Outcomes (LO). Moderate standard deviations and balanced mean-median distances imply relatively symmetric distributions and adequate variability, further supporting the quality of data used in reflective modeling (Hair et al., 2019).

Based on data given in Table 2, AI Interaction has the highest mean scores (around 3.96 - 4.21) with slightly lower standard deviations. This shows high central tendency as well as consistent response patterns. This result indicates that experience with AI systems is not only perceived well, but also constant among the respondents, which qualifies it as an important substance in the Model under discussion.

After the descriptive analysis, the measurement model is evaluated on the PLS-SEM framework to determine the reliability of the indicators as well as construct validity, and then test the structural relations.

4.2. Measurement Model Assessment

4.2.1. Indicator Reliability

Table 3. Outer loadings of measurement items

Sources: The Author (2026)

	AI	LO	MK	MM	SE
LO1		0.763			
LO2		0.730			
LO3		0.786			
LO4		0.755			
LO5		0.770			
MK1			0.662		
MK2			0.736		
MK3			0.723		
MK4			0.688		
MK5			0.657		
MK6			0.747		
MK7			0.746		
MK8			0.718		
MK9			0.738		
MM1				0.669	
MM2				0.664	

	AI	LO	MK	MM	SE
MM3				0.733	
MM4				0.708	
MM5				0.709	
MM6				0.766	
MM7				0.725	
MM8				0.649	
MM9				0.762	
SE1					0.640
SE2					0.703
SE3					0.762
SE4					0.766
SE5					0.634
SE6					0.684
SE7					0.709
SE8					0.713
SE9					0.755
AI1	0.717				
AI10	0.741				
AI11	0.686				
AI2	0.673				
AI3	0.664				
AI4	0.776				
AI5	0.730				
AI6	0.706				
AI7	0.669				
AI8	0.748				
AI9	0.748				

The reliability of the indicators was determined by means of standardized outer loading to establish the degree to which each indicator is related to the constructs. Most of the indicators, as indicated in Table 3, showed an outer loading greater than the recommended value of 0.70. This establishes the fact that the latent constructs account for more than half of the variation in their indicators. Even though a few items exhibiting loading that were a bit below this value were retained, they were of great theoretical significance to the constructs of self-efficacy and metacognitive knowledge, and they did not affect the overall construct reliability.

4.2.2. Internal Consistency, Reliability, and Convergent Validity

Internal consistency reliability was assessed using Consistent Reliability, while Average Variance Extracted (AVE) was used to evaluate convergent validity.

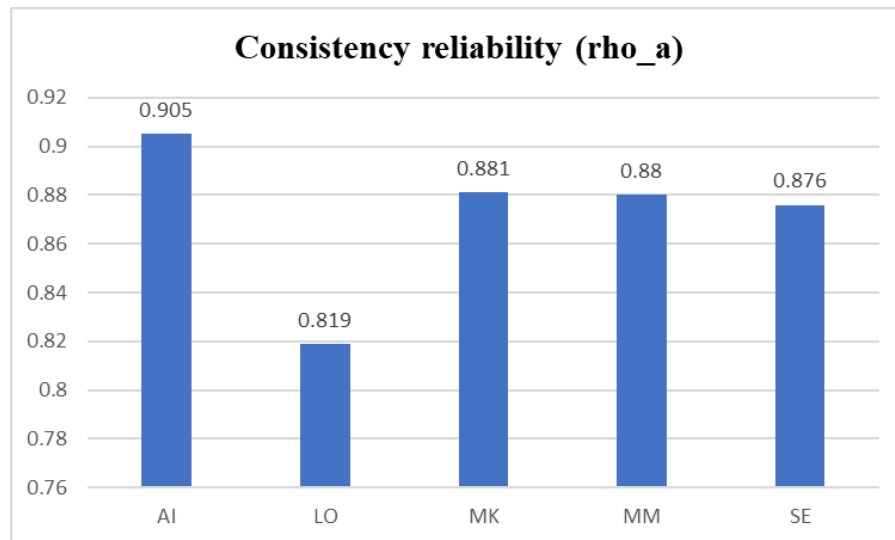


Figure 6. Consistent reliability values of the measurement model

Source: The Author (2026)

The results in Figure 6 indicate that all constructs demonstrate satisfactory internal consistency reliability (rho_a) since it is often recommended in PLS-SEM as it provides a more precise reliability estimate compared to Cronbach's alpha and traditional Composite Reliability (Hair et al., 2019). To be more detailed, Dijkstra-Henseler's consistency reliability (rho_a) values range from 0.819 to 0.905, exceeding the recommended threshold of 0.70. These findings suggest that the indicators consistently measure their respective latent constructs.

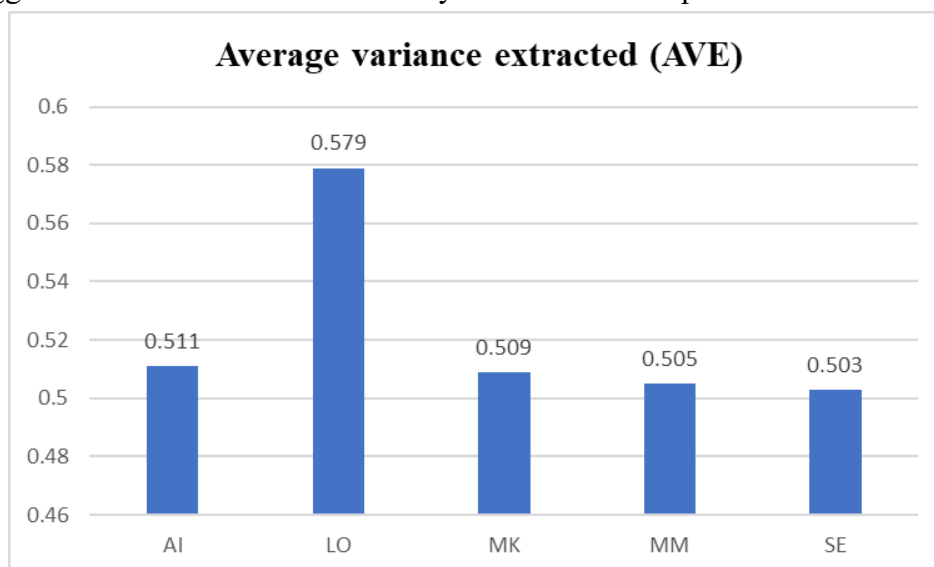


Figure 7. Average variance extracted (AVE) values of the measurement model

Source: The Author (2026)

The results of convergent validity are also supported by the results of the Average Variance Extracted (AVE), where all constructs can be seen to report AVEs that are well above the minimum criterion of 0.50, as shown in Figure 7. It means that both constructs can account for the greater part of the variance of indicators, and there is sufficient convergence. The AVE of Learning Outcomes (LO) is relatively high (0.579), which is why this construct is specifically

well represented by the indicators, and the other constructs also have a high level of methodological consistency (Fornell and Larcker, 1981; Hair et al., 2021).

4.2.3. Discriminant Validity

Discriminant validity was examined using the Heterotrait–Monotrait ratio (HTMT), which has been recommended as a more rigorous criterion than traditional methods such as the Fornell Larcker criterion (Henseler et al., 2015).

Table 4. Heterotrait-Monotrait Ratio of Correlations

Source: The Author (2026)

	AI	LO	MK	MM	SE
AI					
LO	0.864				
MK	0.679	0.825			
MM	0.700	0.733	0.819		
SE	0.560	0.691	0.740	0.671	

Table 4 shows the Heterotrait-Monotrait ratio of the correlations between all the construct pairs, and this is the last and the most stringent criterion to evaluate the discriminant validity. The findings are meaningful: the ratios are mostly below the conservative figure of 0.85. To be more specific, although the HTMT value for AI Interaction (0.864) marginally exceeds the conservative 0.85 benchmark, it remains well within the acceptable upper limit of 0.90 (Henseler et al., 2015), thereby preserving discriminant validity. These findings suggest that Self-Efficacy (SE), Metacognition Knowledge (MK), Metacognition Monitoring (MM), and AI Interaction (AI) are unique entities. It establishes the fact that the respondents are able to distinguish between their personal abilities and their engagement with the use of AI tools, which proves the structural soundness of the given model.

4.3. Summary of Measurement Model Findings

In summary, the measurement model was found to possess strong psychometric characteristics. The constructs all meet the suggested levels of internal consistency reliability, which are represented by Dijkstra-Henseler's rho values of more than 0.70. Convergent validity is observed by having the Average Variance Extracted (AVE) values being above the minimum value of 0.50, which is the acceptable minimum aspect in ensuring that the indicators effectively measure their respective underlying latent constructs.

Besides, the Heterotrait-Monotrait Ratio (HTMT) results also argue in favor of discriminant validity, and the almost inter-construct correlations are less than the conservative value of 0.85. It was important to note that AI Interaction proves to be a valid, empirically discrete construct, characterized by high internal consistency and distinct discriminant validity with respect to cognitive, metacognitive, and outcome-related constructs. All these results hold to prove that the measurement tool is methodologically sound and can be used to study the contributions of both self-efficacy and AI interaction into the learning scenarios of higher education settings.

4.4. Positioning of the Results within the Research Scope

It is worth pointing out that, since this working paper is rather exploratory, the analysis is confined solely on measurement validation. This meant that at this point, the structural relationships, which are the direct and mediating effects between the latent constructs, were not statistically tested. The main input of this section is that a good and sound measurement model is developed. This is one of the mandatory conditions, which will guarantee that the structural models analysis in the future will be conducted on the grounds of the correct and consistent depiction of the latent constructs.

5. Conclusion

Several limitations should temper interpretation. First, although the study targeted 400 responses for power, the final analytic sample was 302; this may be adequate for exploratory PLS-SEM measurement validation, but stronger structural inference - especially sequential mediation - would benefit from larger samples and robustness checks. Second, generalizability is limited to undergraduates in Ho Chi Minh City who actively use GenAI chatbots for academic purposes; replication across regions, disciplines, and assessment cultures is needed, given differences in AI-use norms and academic integrity policies. Third, self-report outcomes may overstate learning because fluency and convenience can inflate perceived gains; future work should triangulate with objective performance and behavioral traces, including calibration measures linking confidence to correctness (Schraw, 2009). Finally, replicability would improve by clarifying instrument documentation, including item selection, adaptation, and any shortening procedures.

Yet a provocative question remains: could strategically omitting GenAI interaction - at least at key moments of planning, self-testing, and verification - yield deeper metacognitive engagement and more accurate calibration? This working paper motivates the next phase of research. Building on this validated instrument set, the forthcoming full scientific study will test the structural pathways with bootstrapped indirect effects, incorporate objective performance and calibration measures, and compare learning conditions with and without AI assistance to identify when AI support strengthens learning versus when it attenuates metacognitive agency (Bućinca et al., 2021; Fan et al., 2024; Klingbeil et al., 2024). Ultimately, this agenda shifts the field from optimizing AI-assisted outputs to safeguarding human intellectual agency in an era where GenAI prevails.

TÀI LIỆU THAM KHẢO

- A, C. S., & Neuman, G. (2008). A model of metacognition, achievement goal orientation, learning style and self-efficacy. *Learning Environments Research*, 11, 131–151.
- Richmond, V. P., Corham, J., & McCroskey, J. C. (1987). The relationship between selected immediacy behaviors and cognitive learning. In M. McLaughlin (Ed.), *Communication yearbook 10* (pp. 574-590). Beverly Hills, CA: Sage.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. W.H. Freeman.
- Bandura, A. (2001). Social Cognitive theory: an Agentic Perspective. *Annual Review of Psychology*, 52(1), 1–26.
- Brown A. L. (1987). Metacognition, executive control, self-regulation, and other more mysterious mechanisms. In *Metacognition, motivation, and understanding* (pp. 65–116). Lawrence Erlbaum Associates.
- Buçınca, Z., Maja Barbara Malaya, & Gajos, K. Z. (2021). To Trust or to Think: Cognitive Forcing Functions Can Reduce Overreliance on AI in AI-assisted Decision-making. *ResearchGate*.
- Dignath, C., Ewijk, van, Perels, F., & Fabriz, S. (2023). Let learners monitor the learning content and their learning behavior! A meta-analysis on the effectiveness of tools to foster monitoring. *Educational Psychology Review*, 35, 62.
- Fernandes, D., Villa, S., Nicholls, S., Haavisto, O., Buschek, D., Schmidt, A., Kosch, T., Shen, C., & Welsch, R. (2026). AI makes you smarter but none the wiser: The disconnect between performance and metacognition. *Computers in Human Behavior*, 175, 108779.
- Flavell, J. H. (1979). Metacognition and Cognitive Monitoring: a New Area of Cognitive-Developmental Inquiry. *American Psychologist*, 34(10), 906–911.
- Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, 31(1), 2–24.
- Halmo, S. M., Yamini, K. A., & Stanton, J. D. (2024). Metacognition and Self-Efficacy in Action: How First-Year Students Monitor and Use Self-Coaching to Move Past Metacognitive Discomfort During Problem Solving. *CBE- Life Sciences Education*, 23(2).
- Hayat, A. A., & Shateri, K. (2019). The Role of Academic Self-Efficacy in Improving Students' Metacognitive Learning Strategies. *Journal of Advances in Medical Education & Professionalism*, 7(4), 205–212.
- 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.
- J. H. Flavell. (1987). Speculations about the nature and development of metacognition. In *Metacognition, motivation, and understanding* (pp. 21–29). Lawrence Erlbaum Associates.
- Klingbeil, A., Grützner, C., & Schreck, P. (2024). Trust and Reliance on AI - an Experimental Study on the Extent and Costs of Overreliance on AI. *Computers in Human Behavior*, 160(108352), 108352–108352.
- Pintrich, P. R. (1999). The role of motivation in promoting and sustaining self-regulated learning. *International Journal of Educational Research*, 31(6), 459–470.

Rahimi, M., & Abedi, S. (2014). The relationship between listening self-efficacy and metacognitive awareness of listening strategies. *Procedia - Social and Behavioral Sciences*, 98, 1454–1460.

Schraw, G. (2009). Measuring metacognitive judgments. In J. Dunlosky & A. C. Graesser (Eds.), *Handbook of metacognition in education*.

Schraw, G., Crippen, K. J., & Hartley, K. (2006). Promoting Self-Regulation in Science Education: Metacognition as Part of a Broader Perspective on Learning. *Research in Science Education*, 36(1-2), 111–139.

Schraw, G., & Dennison, R. S. (1994). Assessing Metacognitive Awareness. *Contemporary Educational Psychology*, 19(4), 460–475.

Schraw, G., & Moshman, D. (1995). Metacognitive theories. *Educational Psychology Review*, 7(4), 351–371.

Schunk, D. H., & Zimmerman, B. J. (2008). *Motivation and self-regulated learning: theory, research, and applications*. Routledge.

Yokoyama, S. (2024). Impact of academic self-efficacy on online learning outcomes: a recent literature review. *PubMed*, 23, 960–966.

Zhai, Y., & Behzad Nezakatgoo. (2025). Evaluating AI-Powered Applications for Enhancing Undergraduate Students' Metacognitive Strategies, Self-Determined Motivation, and Social Learning in English Language Education. *Scientific Reports*, 15(1).

Zhao, X., Lynch, J. G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. 37, 197–206. https://www.researchgate.net/publication/228424966_Reconsidering_Baron_and_Kenny_Myths_and_truths_about_mediation_analysis

Zimmerman, B. J. (2000). Attaining Self-Regulation. *Handbook of Self-Regulation*, 13–39.

Zimmerman, B. J. (2008). Investigating Self-Regulation and Motivation: Historical Background, Methodological Developments, and Future Prospects. *American Educational Research Journal*, 45(1), 166–183.