**International Journal of Evaluation and Research in Education (IJERE)**
Vol. 15, No. 5, October 2026, pp. 4245~4255
ISSN: 2252-8822, DOI: 10.11591/ijere.v15i5.39649 

# Bringing the flipped classroom to Marxist-Leninist philosophy: evidence from teaching Indian medical students in Vietnam

**Phuoc Tai Nguyen[1], Thanh Hung Nguyen[1], Thuan Quy Nguyen[2]**
[1]Faculty of Basic Sciences, Can Tho University of Medicine and Pharmacy, Can Tho, Vietnam
[2]Faculty of Culture, Tourism and Social Work, Dong Thap University, Can Tho, Vietnam

**Article Info**

***Article history:***

Received Mar 22, 2026
Revised Jul 9, 2026
Accepted Aug 6, 2026

***Keywords:***

Critical thinking mediation
Digital learning technologies
Flipped classroom
International medical students
Political theory education

**ABSTRACT**

This study tests whether a flipped classroom model supported by digital technologies improves learning outcomes in a compulsory Marxist-Leninist philosophy course taught in English to Indian medical students at Can Tho University of Medicine and Pharmacy (CTUMP), Vietnam, and, more importantly, why it does so. Flipped learning has been evaluated extensively in science, engineering, and language education, but almost never in political theory, and rarely with international, non-native-speaking cohorts enrolled in compulsory ideological coursework. Most evaluations also assume, rather than measure, the psychological pathway linking digital preparation to higher-order thinking. A convergent mixed-methods design assigned 420 Indian students to an experimental group (EG) (n=210) taught through a learning management system (LMS), micro-learning videos, pre-class formative quizzes, and interaction-intensive class sessions, or to a control group (CG) (n=210) receiving lecture-based instruction. Validated scales measured digital learning access (DLA), cognitive load barriers (CLB), self-regulated learning (SRL), classroom interaction (CLA), critical thinking ability (CTA), and technology/learning analytics acceptance (TLA). Independent-samples t-tests favored the EG on every construct (d=.96-1.74), and bootstrapped mediation analysis showed that CLA carried 55.8% of the effect of DLA on critical thinking, with the largest gains among students entering with lower grade point averages (GPA). CLA, and not digital exposure by itself, is the principal mechanism through which technology-enhanced preparation raises higher-order thinking. The model offers an evidence-based and equity-enhancing route for political theory instruction in international medical education.



***Corresponding Author:***

Thanh Hung Nguyen
Faculty of Basic Sciences, Can Tho University of Medicine and Pharmacy
179 Nguyen Van Cu Street, An Khanh Ward, Can Tho, Vietnam
Email: nthung@ctump.edu.vn

## 1. INTRODUCTION

Universities across Europe and Asia are under pressure to move teaching onto digital platforms, and the way content is delivered and assessed has changed accordingly [1], [2]. Political theory courses, and Marxist-Leninist philosophy in particular, face a version of this pressure that other disciplines do not: instructors have to modernize delivery while preserving the ideological rigor the subject demands. The difficulty grows in international medical education, where the same instructors bridge linguistic and cultural distance while remaining accountable to compulsory curricular standards, a tension recent reviews of digital-technology adoption describe in similar terms [3], [4].

At Can Tho University of Medicine and Pharmacy (CTUMP), Marxist-Leninist philosophy is compulsory for every degree program, including the English-medium tracks taken by international students. Indian students, the largest such cohort, arrive with strong debating traditions yet must process abstract philosophical content in a second language while managing demanding clinical schedules. Our earlier study of this cohort's intercultural adjustment found that structured, technology-supported teaching improved their sense of inclusion in Vietnamese political theory classrooms [5], which prompted the present attempt to identify the mechanism behind that gain. Under conventional lectures these students often fall back on surface learning, reproducing definitions rather than arguing with them [6].

The flipped classroom offers a constructivist alternative [7], [8]. Remembering and understanding move to a pre-class phase carried by micro-learning videos, structured readings, and formative quizzes, while class time is reserved for analysis, discussion, and problem-solving [8], [9]. Cognitive load is redistributed across the learning cycle rather than compressed into one session, and for students facing linguistic and conceptual barriers at once, self-paced preparation lets language processing happen before collaborative work begins. The redistribution rests on a claim that time freed from transmission is reinvested in the interaction that produces higher-order thinking.

Meta-analyses support flipped classroom effectiveness across disciplines [8], [10], but three gaps persist. Almost all existing studies sit in science, engineering, or language learning [11], [12], leaving political theory and the humanities untested, and very little research has examined the model with international, non-native-speaking students in compulsory ideological coursework. Most evaluations also rest on pre-post score comparisons, so the psychological mechanism connecting digital preparation to higher-order gains remains asserted rather than tested [13]. This study addresses all three.

To the authors' knowledge, this is the first controlled evaluation of a flipped classroom model in a compulsory Marxist-Leninist philosophy course delivered to an international, non-native-speaking medical cohort, and the first study in this content area to treat classroom interaction (CLA) as a statistically measured mediator rather than an assumed one. It extends our related work on multicultural inclusion in the same teaching context [5] by isolating the pathway through which technology-supported flipped instruction is associated with higher-order cognitive outcomes, and by asking whether the benefits are distributed equitably across prior academic performance and gender. The disciplinary extension into political theory and the mechanistic test of the interaction pathway together constitutes the paper's original contribution.

The study reports a quasi-experimental investigation at CTUMP in which a flipped model, supported by a learning management system (LMS), micro-learning videos, and an artificial intelligence (AI)-assisted knowledge retrieval tool, was implemented for a Marxist-Leninist philosophy course. It tests whether the approach improves six theoretically grounded constructs, and whether CLA mediates the link between digital learning resources and critical thinking. Because assignment was not randomized, causal language is hedged throughout.

The findings speak to pedagogical innovation in political theory, technology-enhanced learning in medical education, and international student pedagogy. Section 2 reviews the theory and evidence behind the hypotheses and section 3 sets out the method. Section 4 reports the results, section 5 discusses them and states the limitations, and section 6 concludes.

## 2. LITERATURE REVIEW

### 2.1. Theoretical framework

Three frameworks converge on one testable proposition: digital pre-class resources improve critical thinking mainly by enabling CLA rather than substituting for it. Bloom's revised taxonomy supplies the structural rationale, separating remembering and understanding, which move to the pre-class phase, from application, analysis, evaluation, and creation, which are given priority in class [9], [14]. Face-to-face time is thereby concentrated on tasks that gain most from peer interaction and expert guidance.

Vygotsky's social constructivism frames in-class collaboration as the primary mechanism of knowledge construction [15]. Within the zone of proximal development, peer discussion and instructor scaffolding carry learners to an understanding they could not reach alone, which matters for students meeting unfamiliar philosophical frameworks for the first time. The theory predicts that any benefit of pre-class preparation should be realized chiefly through the quality of the classroom dialogue that follows, which is the pathway modelled in subsection 3.5.

Cognitive load theory separates intrinsic load, the inherent difficulty of the content, from extraneous load imposed by poor instructional design and germane load devoted to schema formation [14]. Micro-learning videos, short and topic-focused and subtitled in English, are designed to cut extraneous load and free capacity for deeper processing in class. The theory supplies the mechanism by which preparation translates into more productive engagement, complementing the social-constructivist account rather than competing with it.

Taken together, the three frameworks predict a chain, not three independent effects. Reduced load during preparation frees capacity for better CLA, and that interaction, rather than digital exposure itself, drives gains on Bloom's higher-order categories. The effect of digital learning access (DLA) on critical thinking should therefore be transmitted substantially, though not wholly, through CLA, which section 4.6 tests by mediation analysis.

### 2.2. Flipped classroom in higher education

Meta-analytic evidence shows modest but reliable gains. Alten *et al.* [8] reported g=.36 across 114 studies, and Låg and Sæle [16] found g=.24 across 272 samples after correcting for publication bias. The effects hold across disciplines and populations, though they are systematically larger where flipped instruction is paired with active, peer-based in-class work, a moderating pattern already visible in Bishop and Verleger's survey [7].

In medical education the model has improved clinical reasoning, retention, and engagement. Hew and Lo [17] meta-analysis found consistently positive if heterogeneous effects on knowledge outcomes across medicine, nursing, and pharmacy, and Spaic *et al.* [18] synthesizing 141 studies and more than 21,000 participants, confirmed better knowledge-assessment scores and higher satisfaction. Tolks *et al.* [19] found greater self-study behavior and peer collaboration among flipped medical students than among controls. Effect sizes vary by discipline and outcome, which is one reason context-specific evaluations remain useful.

### 2.3. Digital technologies and learning management systems

LMS are the organizational backbone of digitally enhanced instruction, carrying asynchronous delivery, activity sequencing, progress monitoring, and feedback loops [20]. Video completion rates, quiz performance, and engagement metrics tell instructors and students how preparation converts into classroom readiness [21]. In technology acceptance model (TAM) and its extension unified theory of acceptance and use of technology (UTAUT), perceived usefulness and ease of use are the primary determinants of willingness to engage with digital tools [22], which section 3.4 operationalizes.

AI-powered knowledge retrieval, specifically retrieval-augmented generation (RAG), is an emerging way of protecting content accuracy in ideologically sensitive disciplines. Unlike general-purpose large language models, RAG systems constrain responses to a pre-validated knowledge base, which lowers the risk of factual distortion [23], [24]. Conceptual precision and alignment with official standards are non-negotiable in Marxist-Leninist philosophy. Here the RAG component was a supplementary post-class resource, not an object of formal outcome measurement, a scope limitation acknowledged in section 5.6.

### 2.4. Flipped classroom for international students in political theory

Work at the intersection of flipped instruction, international students, and political theory is sparse. Studies of international students in general higher-education settings indicate that structured preparatory resources and low-stakes formative assessment ease the transition into unfamiliar academic and linguistic environments [25]. Our related work on Indian students' intercultural experience in this same setting found that technology-supported, interactionally rich instruction went together with a stronger sense of belonging and less communication anxiety [5], a finding the present study extends by testing the cognitive mechanism underlying it.

Existing work on Marxist-Leninist philosophy pedagogy in Vietnamese universities has proposed frameworks without testing them under controlled designs [26], [27]. This study is, to our knowledge, the first controlled evaluation of flipped instruction in a compulsory political-theory course taught to an international medical cohort, and the first to test CLA as a measured mediator in this content area. It answers calls for empirical rather than conceptual evaluation of technology-enhanced ideological education [28].

## 3. METHOD

### 3.1. Research design

The study used a convergent mixed-methods design with a quasi-experimental component for quantitative comparison and an exploratory qualitative component for interpretation. A non-equivalent control group (CG) structure assigned 420 Indian students enrolled in the Marxist-Leninist philosophy course to an experimental group (EG) (n=210) or a CG (n=210) on the basis of class scheduling, with the groups matched on academic year, gender distribution, and prior grade point average (GPA). Because assignment followed administrative sections rather than randomization, selection effects cannot be ruled out, and the constraint is treated as a limitation throughout and addressed directly in section 5.6.

### 3.2. Participants

Participants were Indian nationals enrolled in the doctor of medicine program at CTUMP. The study targeted Cohort 51 (K51), the entire eligible population registered for the Marxist-Leninist philosophy module in the Spring 2026 semester. Complete sampling was used, and after five incomplete responses were excluded, the analytic sample was N=420, as described in Table 1.

Table 1. Demographic characteristics of participants (N=420)

| Characteristic | EG (N=210) | CG (N=210) | Total (N=420) |
|---|---|---|---|
| Gender: male | 112 (53.3%) | 108 (51.4%) | 220 (52.4%) |
| Gender: female | 98 (46.7%) | 102 (48.6%) | 200 (47.6%) |
| Academic year 1 | 121 (57.6%) | 123 (58.6%) | 244 (58.1%) |
| Academic year 3 | 89 (42.4%) | 87 (41.4%) | 176 (41.9%) |
| Mean GPA (SD) | 3.21 (.44) | 3.18 (.47) | 3.20 (.45) |
| Age: mean (SD) | 20.3 (1.2) | 20.5 (1.3) | 20.4 (1.2) |

### 3.3. Instructional interventions

The experimental condition ran the flipped model across three phases. Before class, students used micro-learning videos of five to seven minutes, structured digital handouts, and concept-mapping resources delivered through the university LMS; each video covered a single conceptual unit and carried English subtitles prepared jointly by content experts and language support staff. Students had to complete a formative quiz before each session, and the LMS recorded views, completion rates, quiz scores, and time on task.

In class, instructors ran Socratic dialogue, case-based discussions linking Marxist-Leninist concepts to medical ethics scenarios, small-group problem-solving, and structured debates on contemporary socio-political questions. These activities required students to apply, analyze, and evaluate philosophical constructs in applied contexts, corresponding to the upper categories of Bloom's revised taxonomy [9]. Instructors were trained to facilitate rather than lecture during this phase.

After class, students completed reflective synthesis tasks and applied assignments, and could consult an AI-assisted retrieval chatbot built on RAG architecture. The system answered only from a pre-validated knowledge base of official textbooks, approved course materials, and policy documents, which preserved ideological accuracy while offering personalized explanation. Use of the chatbot was voluntary and unassessed, and its usage was not analyzed as a separate outcome.

The CG received lecture-based instruction delivered through slides and illustrative examples. Digital course materials were available on the LMS for reference, but no pre-class learning requirement was imposed and in-class activity remained largely receptive. Both arms therefore covered the same content across the same contact hours, differing in the sequencing of cognitive activity rather than in curricular coverage.

The same instructor team taught all sections in both arms, working from a shared lesson plan with weekly coordination meetings and a one-day workshop on flipped facilitation before the semester. These steps limit, without eliminating, the possibility that group differences reflect instructor-specific factors. A structured observation checklist applied during three randomly selected sessions per arm verified that the intended activities were delivered.

### 3.4. Instrumentation

The primary instrument was a 30-item questionnaire written in English and grounded in four validated sources: flipped-classroom research instruments [7], [8], Zimmerman [29] self-regulated learning (SRL) framework, cognitive-load assessment adapted from the NASA-task load index (TLX) [30], and TAM/UTAUT [22]. Items were rated on a 5-point Likert scale and organized into six latent constructs, summarized with sample items and reliability coefficients, as seen in Table 2.

Table 2. Measurement constructs, items, and reliability coefficients

| Code | Construct | Items (n) | α | Sample item |
|---|---|---|---|---|
| DLA | Digital learning access | 4 | .847 | The LMS is clearly organized, making pre-class preparation convenient. |
| CLB | Cognitive load barriers | 4 | .782 | I find it difficult to identify core content when preparing independently. |
| SRL | Self-regulated learning | 4 | .831 | I set learning goals and plan my study time before each class session. |
| CLA | Classroom interaction | 5 | .893 | Group discussions help me understand the substance of philosophical problems. |
| CTA | Critical thinking ability | 4 | .862 | I am more confident asking 'why' and 'how' questions about social and medical issues. |
| TLA | Technology/learning analytics acceptance | 3 | .815 | I believe that combined (technology+traditional) assessments better reflect my abilities. |

Content validity was established through expert review by three faculty members specializing in political theory and one educational measurement specialist. Pilot testing with 45 students, excluded from the main analysis, produced Cronbach's alpha between .782 and .893. Two items flagged as ambiguous by pilot respondents were reworded before the finalized instrument went to the main sample.

### 3.5. Data analysis

Quantitative data were processed in IBM SPSS Statistics 26.0. Independent-samples t-tests compared construct means between groups with Cohen's d as the effect size, and multiple regression identified the strongest predictors of critical thinking. Mediation was tested with the PROCESS macro (model 4), 5,000 bootstrap iterations, and 95% confidence intervals, estimating the indirect effect of DLA on critical thinking through CLA.

Before mediation testing the data were screened for multicollinearity (all variance inflation factors below 2.0), normality of residuals, and homoscedasticity; nothing was detected that would compromise the bootstrap estimates. Baseline equivalence on GPA, gender, and academic year was confirmed, though equivalence on observed variables does not rule out unmeasured confounds such as motivation. Qualitative data from open-ended items and instructor observations were analyzed thematically to triangulate the quantitative findings.

## 4. RESULTS

### 4.1. Construct descriptive statistics and group comparisons

Table 3 presents descriptive statistics and independent-samples t-test results for all six constructs. Five show statistically significant differences in the expected direction. CLB require separate interpretation because lower scores, not higher ones, are favorable for that construct.

The EG scored higher on DLA (d=1.52), SRL (d=1.39), CLA (d=1.74), CTA (d=1.68), and TLA (d=.97), all at p<.001, and every effect size exceeds Cohen's threshold for a large effect. These differences are larger than flipped-classroom meta-analyses typically report [8], [16]. Given the quasi-experimental design, the gap most plausibly reflects the combined effect of the LMS, micro-learning video, and AI-assisted retrieval acting together, and should be read as an upper-bound estimate for this multi-component implementation rather than a baseline expectation for flipped instruction generally.

For CLB the EG reported lower scores (M=2.81, SD=.67 vs. M=3.48, SD=.72; t=10.49, p<.001; d=.96), indicating reduced cognitive barriers under the flipped condition. This supports the claim that micro-learning video design lowers extraneous load by letting students regulate their own information-processing pace [31]. LMS behavioral data corroborate the self-reports: mean video completion was 87.3%, and students rewatched videos an average of 1.8 times per segment.

Table 3. Construct mean scores, standard deviations, and t-test results by group

| Code | Construct | EG M (SD) | CG M (SD) | t | p | d |
|---|---|---|---|---|---|---|
| DLA | Digital learning access | 4.12 (.53) | 3.24 (.61) | 16.94 | <.001 | 1.52 |
| CLB | Cognitive load barriers | 2.81 (.67) | 3.48 (.72) | 10.49 | <.001 | .96 |
| SRL | Self-regulated learning | 4.03 (.58) | 3.19 (.63) | 15.18 | <.001 | 1.39 |
| CLA | Classroom interaction | 4.21 (.51) | 3.14 (.68) | 19.41 | <.001 | 1.74 |
| CTA | Critical thinking ability | 4.09 (.55) | 3.07 (.64) | 18.72 | <.001 | 1.68 |
| TLA | Technology/learning analytics acceptance | 3.96 (.62) | 3.31 (.71) | 10.74 | <.001 | .97 |

Note: d=Cohen's d effect size; all CLB comparisons indicate lower scores are more favorable for the EG

### 4.2. Digital learning resources and learning management system engagement

LMS logs show that EG students engaged substantially with the pre-class materials, with an 87.3% completion rate well above typical benchmarks for asynchronous environments and an average of 1.8 replay cycles per segment, indicating deliberate return to difficult content [19]. Open-ended responses reinforce the pattern: students cited pausing, rewinding, and replaying as what made preparation possible, particularly for philosophical terminology. Several described lower anxieties about speaking in class because they had already processed the content at their own pace.

### 4.3. Self-regulated learning development

The large SRL effect (d=1.39) points to a shift in study habits rather than incremental improvement. EG students reported more strategic goal-setting, better time management, and more metacognitive monitoring, the core components of SRL as Zimmerman [29] defines them. The mandatory preparation structure appears to work as an external scaffold for these strategies, consistent with Zimmerman's cyclical model of forethought, performance, and self-reflection.

About a third of EG students nevertheless reported difficulty balancing pre-class video viewing against clinical placement schedules. This points to an actionable revision: shorter videos, segmented more finely, to fit the scheduling constraints particular to medical students. It also shows that self-regulation gains, large on average, were uneven across the sample.

### 4.4. Classroom interaction and quality of discourse

CLA produced the largest effect of the six (d=1.74), so reallocating class time from listening to discussion coincided with the biggest behavioral change measured here. EG students reported more participation, more frequent volunteering of opinions, and greater psychological safety in expressing dissenting views, and instructor observations agree: speaking turns were longer and more evenly distributed than in the instructor-led control sections. Socio-medical case studies linking Marxist-Leninist concepts to bioethics, health equity, and the social determinants of health worked particularly well, since Indian students, whose prior education cultivates argumentation, responded strongly to structured debate about concrete professional dilemmas. Whether the mechanism transfers to courses whose students lack comparable debate traditions is an open question this study cannot settle.

### 4.5. Critical thinking development

Critical thinking scores in the EG exceeded those in the CG (d=1.68), with the clearest improvements in analytical reasoning, evaluative judgment, and the application of philosophical concepts to medical contexts. Experimental students reported more confidence in defending a position and more frequent questioning of underlying assumptions, behaviors characteristic of Bloom's upper taxonomy levels [7], and the gains appeared in both academic years. Open-ended responses reveal a parallel shift: CG students more often described learning as memorization, whereas experimental students described the course as a place to develop analytical tools for clinical reasoning. Since both sources are self-report, this contrast corroborates the quantitative pattern without independently proving it.

### 4.6. Mediation analysis: the role of classroom interaction

The mediation analysis in Table 4 tested whether CLA mediates the relationship between DLA and CTA, the central proposition of section 2.1. It is the study's principal empirical contribution because it moves past group comparison to test the chain implied by Bloom's taxonomy, social constructivism, and cognitive load theory. Path coefficients are read against the theoretical predictions.

The indirect effect of DLA on CTA through CLA was b=.35 (95% CI [.28, .43]), accounting for 55.8% of the total effect. The direct effect remained significant (b=.27, p<.001) but was substantially reduced, so the mediation is partial rather than full. The pattern matches the proposition under test: digital resources are associated with critical thinking gains mainly because they enable higher-quality CLA, not because they deliver content directly. Because the design is quasi-experimental and confounds such as motivation and baseline engagement went unmeasured, this is evidence consistent with a mediating mechanism rather than proof of a causal pathway. It aligns with social constructivist principles [15] and with research holding that the educational value of technology is realized through the pedagogical structures it enables [24].

Table 4. Mediation analysis results: DLA→CLA→CTA (N=420)

| Path | β | SE | t | p | 95% CI |
|---|---|---|---|---|---|
| DLA→CLA (path a) | .61 | .04 | 14.27 | <.001 | [.53, .69] |
| CLA→CTA (path b) | .58 | .05 | 12.04 | <.001 | [.48, .68] |
| DLA→CTA (direct path c') | .27 | .06 | 4.85 | <.001 | [.16, .38] |
| Indirect effect (a×b) | .35 | .04 | - | - | [.28, .43]* |
| Proportion of total effect mediated | 55.8% | - | - | - | - |

Note: Bootstrap CIs based on 5,000 iterations; *CI does not include zero, confirming significant mediation

### 4.7. Equity implications

Sub-group analysis by prior performance showed that students with pre-study GPAs below 3.0 gained most on both SRL and CTA relative to their control-group counterparts, so repeated video review, immediate quiz feedback, and structured discussion may matter most to students who struggle with conceptually dense material. Because this result comes from a single institution and was not pre-registered, it should be treated as exploratory. Gender-based analysis found no significant differences, consistent with prior work reporting that active-learning designs do not systematically disadvantage either gender, though the analysis was not powered to detect small interaction effects. The equity-relevant pattern in these data therefore concerns prior academic performance rather than gender.

## 5. DISCUSSION

The flipped classroom model, integrated with digital technologies, went together with large improvements in DLA, SRL, CLA, CTA, and TLA, and with reduced cognitive-load barriers. The contribution is mechanistic rather than comparative: CLA explained 55.8% of the effect of digital learning resources on critical thinking.

### 5.1. The mediating role of classroom interaction: theoretical interpretation

Partial mediation, with CLA carrying just over half the total effect, maps onto Vygotsky's account of learning as social rather than solely individual [15]. On that account digital pre-class materials cannot generate higher-order thinking by themselves; they prepare students to take part in the dialogue through which such thinking is built. The findings qualify the claim as well, since 44.2% of the total effect remained direct: digital exposure retains independent value, so social mediation and individual cognitive processing appear to work as complementary pathways here.

The result converges with cognitive load theory and extends it [31]. CLT predicts that reducing extraneous load during preparation frees germane resources for deeper processing, and the data fit: the EG markedly lower CLB scores (d=.96) accompanied its largest gains on CLA and CTA. CLT alone does not explain why the benefit routes through CLA rather than individual post-class reflection, and the mediation result shows that load reduction and interaction jointly account for critical-thinking gains.

Bloom's revised taxonomy supplies the structural logic for moving lower-order tasks before class [9]. It predicts disproportionate gains in application, analysis, and evaluation, and the large critical-thinking effect (d=1.68) fits that prediction. Most prior flipped-classroom studies report that in-class activity accompanies higher-order gains without testing the pathway [8], [10]; the mediation analysis tests Bloom's implicit causal claim more directly.

### 5.2. Comparison with prior flipped-classroom evidence

The effect sizes reported here (d=.96-1.74) are considerably larger than those from general higher-education meta-analyses, whether Alten *et al.* [8] g=.36 across 114 studies or Låg and Sæle [16] g=.24 across 272 samples after correction for publication bias. They sit closer to the upper range of health-professions syntheses, where Hew and Lo found consistently positive effects varying by discipline and outcome [17] and Spaic *et al.* [18] reported standardized mean differences near .90 across more than 21,000 participants. The most plausible explanation is the multi-component design; dismantling studies would test this directly.

The comparison also marks a boundary condition. Tolks *et al.* [19] treated self-study behavior and peer collaboration as the inverted classroom's principal behavioral signatures in medical education. The present findings reproduce that pattern in a very different discipline and with an international rather than domestic cohort, which suggests the core psychological mechanism of flipped instruction is fairly domain-general, even if the argumentation traditions amplified its magnitude here.

### 5.3. Self-regulated learning and the role of external scaffolding

The large SRL gain (d=1.39) is consistent with Zimmerman model [29], in which self-regulation is a cyclical process of forethought, performance, and self-reflection that can be scaffolded externally before it is internalized. Mandatory pre-class quizzes and video-completion requirements appear to have worked as exactly this kind of scaffold. The residual time-management difficulties reported by roughly a third of experimental students show that external scaffolding does not remove the self-regulatory burden.

This tension between structure and autonomy runs through the self-regulated-learning literature, which has long debated whether imposed deadlines support autonomous regulation or undermine it once the scaffold is withdrawn. The present study cannot settle the question, since it measured SRL within a single semester of mandatory structure. These gains are therefore best read as short-term behavioral change rather than durable self-regulatory competence.

### 5.4. Technology acceptance and the role of AI-assisted resources

The TLA gain (d=.97) is consistent with TAM/UTAUT's prediction that perceived usefulness and ease of use drive engagement with digital tools [22]. Students' comments about pausing, rewinding, and replaying map onto ease of use, and their reports of reduced in-class anxiety map onto perceived usefulness. Survey evidence on RAG-based educational chatbots points the same way: constraining AI responses to a validated knowledge base raises perceived trustworthiness and willingness to use the tool in factually sensitive domains [23], [24], and generative-AI tools are seen as beneficial mainly when their outputs are accurate and pedagogically appropriate [32].

The study did not measure chatbot usage frequency or its independent contribution to critical thinking. The omission is not trivial, since there is growing evidence that over-reliance on conversational AI can accompany reduced independent reasoning rather than enhanced critical thinking [33]. Future iterations

should log chatbot interaction as a distinct variable and draw on AI-literacy frameworks so that students evaluate rather than accept AI-generated explanations [34].

### 5.5. Equity considerations in international medical education

The larger gains among lower-performing students echo a recurring theme in the literature on technology-enhanced learning: structured, repeatable digital resources tend to help most those who struggle with unscaffolded, one-pass instruction [35]. This is an encouraging signal for institutions serving heterogeneous international cohorts, and it fits evidence that AI-driven analytics can identify students at risk before gaps widen [36]. It should not be overstated, because the sub-group finding was exploratory and the equity literature warns that technology can widen achievement gaps as easily as narrow them, depending on implementation quality and on access to devices and connectivity [37], neither of which was formally measured here.

The absence of gender-based differences is reassuring, if modest, and consistent with the general finding that well-designed active-learning environments do not systematically disadvantage either gender. The study was not powered to detect small gender interaction effects, so the null result is an absence of evidence for disadvantage rather than evidence of equivalence [38]. A pre-registered, adequately powered follow-up would allow that claim to be tested with more confidence.

### 5.6. Limitations

Six limitations qualify the causal claims that can be drawn. The quasi-experimental design assigned students by administrative scheduling rather than random allocation, so unmeasured confounds such as motivation or prior exposure to active learning cannot be ruled out, even though the groups were equivalent on GPA, gender, and academic year at baseline. All six constructs were measured with self-report Likert scales, open to social-desirability bias and common-method variance; the LMS behavioral data triangulate the preparation constructs but not CLA or critical thinking. The study ran at a single institution with a single national-origin cohort over one semester, which constrains the generalizability of both the effect sizes and the mediation pathway.

Three further limitations concern the intervention. The AI-assisted RAG chatbot was available to experimental-group students but its use was neither logged nor analyzed, so its contribution cannot be separated from the LMS and video components. The unusually large effect sizes most plausibly reflect the multi-component nature of the intervention, and because the components were bundled rather than manipulated independently, the marginal contribution of any single one cannot be determined. The mediation analysis rests on cross-sectional data from one time point, so it establishes a pattern consistent with the hypothesized chain rather than temporal precedence among constructs.

### 5.7. Implications and future research

Three implications follow for practice. Medical schools teaching political theory to international students should consider flipped adoption where linguistic and cultural barriers compound the difficulty of abstract content [25], [39]. Design matters as much as adoption: micro-learning segmentation, English subtitling, and contextual linkage to medical scenarios are what reduce extraneous load [31]. The larger gains among lower-performing students suggest flipped instruction may narrow achievement gaps in large international cohorts, though this requires confirmation [35]. Future work should pursue randomized designs where feasible, observe across multiple semesters to capture retention of SRL gains, and isolate the contribution of individual components, particularly chatbot interaction patterns and instructor facilitation style, in courses whose students lack the argumentation traditions of this Indian cohort [23].

## 6. CONCLUSION

This study set out to test, rather than assume, the mechanism through which a flipped classroom model integrated with digital technologies improves critical thinking in a compulsory Marxist-Leninist philosophy course for international medical students. CLA, and not digital exposure alone, is the principal pathway linking technology-enhanced preparation to higher-order thinking; DLA, SRL, and TLA also showed large gains relative to lecture-based instruction, and the students who gained most entered with the weakest prior performance. The contribution is both disciplinary, extending flipped-classroom research into political theory, and mechanistic, replacing an assumed pathway with a measured one. Six limitations, chief among them the quasi-experimental design and the bundling of intervention components, mean these results should guide rather than settle practice. Flipped instruction nonetheless offers a workable route to turning ideologically sensitive and conceptually demanding courses into places where critical thinking is built.

**ACKNOWLEDGMENTS**

The authors thank Can Tho University of Medicine and Pharmacy for institutional support, and the 420 Indian students who took part in the survey and gave their time and their responses. The authors also thank those who assisted with survey distribution, data collection, and research coordination; their contributions do not meet the criteria for authorship, but the work could not have been carried out without them. All individuals named here have consented to be acknowledged.

**FUNDING INFORMATION**

The author(s) received no financial support for the research, authorship, and/or publication of this article.

**AUTHOR CONTRIBUTIONS STATEMENT**

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

| Name of Author | C | M | So | Va | Fo | I | R | D | O | E | Vi | Su | P | Fu |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Phuoc Tai Nguyen | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Thanh Hung Nguyen | ✓ | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Thuan Quy Nguyen | ✓ | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ |

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

**CONFLICT OF INTEREST STATEMENT**

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**ETHICAL APPROVAL**

This study involved human participants and was conducted in accordance with ethical standards established by the Vietnam National University system and international research guidelines. Ethical approval was obtained from: Institutional Review Board for Biomedical Research at Can Tho University of Medicine and Pharmacy (Approval No. 25.042.GV/PCT-HĐĐĐ, Date: 26 November 2025). All participants provided informed consent prior to participation. Participants were informed of their right to withdraw at any time without penalty. Data were anonymized and stored securely in password-protected files accessible only to the research team.

**DATA AVAILABILITY**

The data that support the findings of this study are not publicly available because they contain information that could compromise the privacy and confidentiality of student and instructor participants. De-identified, derived datasets supporting the reported analyses are available from the corresponding author, [THN], upon reasonable request and subject to institutional ethical approval.

**REFERENCES**

[1] M. Bond, V. I. Marín, C. Dolch, S. Bedenlier, and O. Zawacki-Richter, "Digital transformation in German higher education: student and teacher perceptions and usage of digital media," *International Journal of Educational Technology in Higher Education*, vol. 15, no. 1, p. 48, Dec. 2018, doi: 10.1186/s41239-018-0130-1.

[2] Y. K. Dwivedi *et al.*, "'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*, vol. 71, p. 102642, Aug. 2023, doi: 10.1016/j.ijinfomgt.2023.102642.

[3] L. Chen, P. Chen, and Z. Lin, "Artificial intelligence in education: a review," *IEEE Access*, vol. 8, pp. 75264–75278, 2020, doi: 10.1109/ACCESS.2020.2988510.

[4] H. Crompton and D. Burke, "Artificial intelligence in higher education: the state of the field," *International Journal of Educational Technology in Higher Education*, vol. 20, no. 1, p. 22, Apr. 2023, doi: 10.1186/s41239-023-00392-8.

[5] N. P. Tai, N. T. Hung, and N. T. Quy, "Flipped classroom technology as a multicultural inclusion strategy for international medical students in Vietnamese political theory education," *Journal for Multicultural Education*, pp. 1–15, Jul. 2026, doi: 10.1108/JME-03-2026-0055.

[6] J. Biggs, D. Kember, and D. Y. P. Leung, "The revised two-factor study process questionnaire: R-SPQ-2F," *The British Journal of Educational Psychology*, vol. 71, no. 1, pp. 133–149, Mar. 2001, doi: 10.1348/000709901158433.

[7] J. L. Bishop and M. A. Verleger, "The flipped classroom: a survey of the research," in *120th ASEE Annual Conference & Exposition*, 2013, pp. 1–18, doi: 10.18260/1-2--22585.

[8] D. C. D. van Alten, C. Phielix, J. Janssen, and L. Kester, "Effects of flipping the classroom on learning outcomes and satisfaction: A meta-analysis," *Educational Research Review*, vol. 28, p. 100281, Nov. 2019, doi: 10.1016/j.edurev.2019.05.003.

[9] D. R. Krathwohl, "A revision of Bloom's taxonomy: an overview," *Theory into Practice*, vol. 41, no. 4, pp. 212–218, Nov. 2002, doi: 10.1207/s15430421tip4104_2.

[10] T. K. F. Chiu, Q. Xia, X. Zhou, C. S. Chai, and M. Cheng, "Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education," *Computers and Education: Artificial Intelligence*, vol. 4, p. 100118, 2023, doi: 10.1016/j.caeai.2022.100118.

[11] K. Cicha, M. Rizun, P. Rutecka, and A. Strzelecki, "COVID-19 and higher education: first-year students' expectations toward distance learning," *Sustainability*, vol. 13, no. 4, p. 1889, Feb. 2021, doi: 10.3390/su13041889.

[12] B. Cope, M. Kalantzis, and D. Searsmith, "Artificial intelligence for education: Knowledge and its assessment in AI-enabled learning ecologies," *Educational Philosophy and Theory*, vol. 53, no. 12, pp. 1229–1245, Oct. 2021, doi: 10.1080/00131857.2020.1728732.

[13] B. Williamson and R. Eynon, "Historical threads, missing links, and future directions in AI in education," *Learning, Media and Technology*, vol. 45, no. 3, pp. 223–235, Jul. 2020, doi: 10.1080/17439884.2020.1798995.

[14] L. W. Anderson and D. R. Krathwohl, *A taxonomy for learning, teaching, and assessing: a revision of Bloom's taxonomy of educational objectives*. New York, NY: Longman, 2001.

[15] L. S. Vygotsky and M. Cole, *Mind in society: development of higher psychological processes*. Cambridge, MA: Harvard University Press, 1978.

[16] T. Låg and R. G. Sæle, "Does the flipped classroom improve student learning and satisfaction? A systematic review and meta-analysis," *AERA Open*, vol. 5, no. 3, pp. 1–17, Jul. 2019, doi: 10.1177/2332858419870489.

[17] K. F. Hew and C. K. Lo, "Flipped classroom improves student learning in health professions education: a meta-analysis," *BMC Medical Education*, vol. 18, no. 1, p. 38, Dec. 2018, doi: 10.1186/s12909-018-1144-z.

[18] D. Spaic *et al.*, "The flipped classroom in medical education: systematic review and meta-analysis," *Journal of Medical Internet Research*, vol. 27, p. e60757, Aug. 2025, doi: 10.2196/60757.

[19] D. Tolks *et al.*, "An introduction to the inverted/flipped classroom model in education and advanced training in medicine and in the healthcare professions," *GMS Zeitschrift fur Medizinische Ausbildung*, vol. 33, no. 3, 2016, doi: 10.3205/zma001045.

[20] G. J. Hwang, H. Xie, B. W. Wah, and D. Gašević, "Vision, challenges, roles and research issues of artificial intelligence in education," *Computers and Education: Artificial Intelligence*, vol. 1, p. 100001, 2020, doi: 10.1016/j.caeai.2020.100001.

[21] E. Lacka, T. C. Wong, and M. Y. Haddoud, "Can digital technologies improve students' efficiency? Exploring the role of virtual learning environment and social media use in higher education," *Computers and Education*, vol. 163, p. 104099, Apr. 2021, doi: 10.1016/j.compedu.2020.104099.

[22] V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis, "User acceptance of information technology: toward a unified view," *MIS Quarterly: Management Information Systems*, vol. 27, no. 3, pp. 425–478, Sep. 2003, doi: 10.2307/30036540.

[23] J. Swacha and M. Gracel, "Retrieval-augmented generation (RAG) chatbots for education: a survey of applications," *Applied Sciences*, vol. 15, no. 8, p. 4234, Apr. 2025, doi: 10.3390/app15084234.

[24] A. V. Y. Lee, S. C. Tan, and C. L. Teo, "Designs and practices using generative AI for sustainable student discourse and knowledge creation," *Smart Learning Environments*, vol. 10, no. 1, p. 59, Nov. 2023, doi: 10.1186/s40561-023-00279-1.

[25] A. Simões, "Preparing students for the global workplace: classroom strategies for developing English and intercultural communication skills," in *ICERI2020 Proceedings*, Nov. 2020, vol. 1, pp. 7962–7971, doi: 10.21125/iceri.2020.1765.

[26] D. T. K. Ng, J. K. L. Leung, S. K. W. Chu, and M. S. Qiao, "Conceptualizing AI literacy: an exploratory review," *Computers and Education: Artificial Intelligence*, vol. 2, p. 100041, 2021, doi: 10.1016/j.caeai.2021.100041.

[27] N. D. Nguyen, "Exploring the role of AI in education," *London Journal of Social Sciences*, no. 6, pp. 84–95, 2023, doi: 10.31039/ljss.2023.6.108.

[28] D. Olcott, X. C. Farran, E. E. G. Echenique, and J. G. Martínez, "Ethics and education in the digital age: global perspectives and strategies for local transformation in Catalonia," *RUSC Universities and Knowledge Society Journal*, vol. 12, no. 2, pp. 59–72, Apr. 2015, doi: 10.7238/rusc.v12i2.2455.

[29] B. J. Zimmerman, "Self-regulated learning and academic achievement: an overview," *Educational Psychologist*, vol. 25, no. 1, pp. 3–17, Jan. 1990, doi: 10.1207/s15326985ep2501_2.

[30] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (task load index): results of empirical and theoretical research," in *Human Mental Workload*, P. A. Hancock and N. Meshkati, Eds., Amsterdam: North-Holland, 1988, pp. 139–183, doi: 10.1016/S0166-4115(08)62386-9.

[31] J. Sweller, "Cognitive load during problem solving: effects on learning," *Cognitive Science*, vol. 12, no. 2, pp. 257–285, Apr. 1988, doi: 10.1207/s15516709cog1202_4.

[32] D. Baidoo-Anu and L. O. Ansah, "Education in the era of generative artificial intelligence (AI): understanding the potential benefits of ChatGPT in promoting teaching and learning," *Journal of AI*, vol. 7, no. 1, pp. 52–62, 2023, doi: 10.61969/jai.1337500.

[33] C. Zhai, S. Wibowo, and L. D. Li, "The effects of over-reliance on AI dialogue systems on students' cognitive abilities: a systematic review," *Smart Learning Environments*, vol. 11, no. 1, p. 28, Jun. 2024, doi: 10.1186/s40561-024-00316-7.

[34] L. K. Allen and P. Kendeou, "ED-AI lit: an interdisciplinary framework for AI literacy in education," *Policy Insights from the Behavioral and Brain Sciences*, vol. 11, no. 1, pp. 3–10, Mar. 2024, doi: 10.1177/23727322231220339.

[35] O. Zawacki-Richter, V. I. Marín, M. Bond, and F. Gouverneur, "Systematic review of research on artificial intelligence applications in higher education – where are the educators?" *International Journal of Educational Technology in Higher Education*, vol. 16, no. 1, p. 39, Dec. 2019, doi: 10.1186/s41239-019-0171-0.

[36] S. Albahli, "Advancing sustainable educational practices through AI-driven prediction of academic outcomes," *Sustainability*, vol. 17, no. 3, p. 1087, Jan. 2025, doi: 10.3390/su17031087.

[37] S. Akgun and C. Greenhow, "Artificial intelligence in education: addressing ethical challenges in K-12 settings," *AI and Ethics*, vol. 2, no. 3, pp. 431–440, Aug. 2022, doi: 10.1007/s43681-021-00096-7.
[38] W. Holmes *et al.*, "Ethics of AI in education: towards a community-wide framework," *International Journal of Artificial Intelligence in Education*, vol. 32, no. 3, pp. 504–526, Sep. 2022, doi: 10.1007/s40593-021-00239-1.
[39] M. St. Fountoulakis, "Evaluating the impact of AI tools on language proficiency and intercultural communication in second language education," *International Journal of Second and Foreign Language Education*, vol. 3, no. 1, pp. 12–26, Oct. 2024, doi: 10.33422/ijsfle.v3i1.768.

## BIOGRAPHIES OF AUTHORS

**Phuoc Tai Nguyen** is a Doctor of Philosophy who graduated from Xiamen University, Fujian Province, China, in 2020, and holds a bachelor's degree in Chinese language. He is currently affiliated with the Faculty of Basic Sciences, Can Tho University of Medicine and Pharmacy, Can Tho, Vietnam. His research interests encompass philosophy, culture, religion, language, and digital transformation in education. He has contributed to various studies on educational philosophy and cultural integration in Vietnam. He can be contacted at email: nptai@ctump.edu.vn.

**Thanh Hung Nguyen** is a lecturer at Can Tho University of Medicine and Pharmacy, Vietnam, with academic interests situated in English language education and applied linguistics. His scholarly work focuses on English-medium instruction (EMI), intercultural communicative competence, language teaching and learning in higher education, and the pedagogical implications of emerging technologies in education. He has contributed to a growing body of research through publications in peer-reviewed journals addressing such topics as international student experiences, EFL teaching practices, intercultural learning, and the integration of digital tools, including ChatGPT, into language education. Through both his teaching and research, he seeks to advance innovative and contextually responsive approaches to language education in Vietnam and beyond. He can be contacted at email: nthung@ctump.edu.vn.

**Thuan Quy Nguyen** holds a Ph.D. in Ethnology (Anthropology) and is currently Head of the Faculty of Culture, Tourism and Social Work at Dong Thap University, Vietnam. Her academic expertise lies at the intersection of cultural studies, education, and language, with a particular focus on folk beliefs, ethnic relations, and religious practices in Southern Vietnam and the Mekong Delta region. She has conducted extensive research on topics such as Vietnamese folk religion, Khmer ethnic relations in border regions, and the cultural significance of religious practices, including studies on Cao Dai architecture and indigenous belief systems. In addition to her research, she has played significant academic and administrative roles within Dong Thap University, contributing to teacher training, curriculum development, and educational innovation at both institutional and regional levels. Her scholarly contributions include numerous publications in national and international journals, as well as participation in ministerial- and provincial-level research projects related to local education development and teacher capacity building. Through her academic work and leadership, she aims to advance culturally grounded and context-responsive approaches to education and social development in Vietnam. She can be contacted at email: ntquy@dthu.edu.vn.