Education Tech Research Dev
https://doi.org/10.1007/s11423-026-10698-9

RESEARCH ARTICLE

AI-enhanced legal pedagogy: a case study of generative AI integration in property law courses

Haifang Liu[1]

Received: 24 February 2026 / Accepted: 5 August 2026
© The Author(s) 2026

Abstract
The rapid emergence of generative artificial intelligence (GenAI) presents transformative opportunities for legal education, yet empirical research examining its pedagogical integration remains limited. This study investigates the impact of GenAI integration on legal reasoning development and learning outcomes in property law courses within Chinese higher education. Grounded in constructivist learning theory, the Technology Acceptance Model (TAM), and the Technological Pedagogical Content Knowledge (TPACK) framework, this research develops an integrated AI-Enhanced Legal Pedagogy Framework to guide effective technology integration in legal education. Employing a mixed-methods case study design, the study engaged 180 law students in a semester-long intervention using iFlytek Spark, combining quantitative assessments through pre-post legal reasoning tests and TAM-based surveys with qualitative data from interviews, reflection journals, and classroom observations. Results indicate significant associations between GenAI-enhanced instruction and improvements in legal reasoning abilities (between-group Cohen's $d = 1.17$), though this effect size warrants cautious interpretation given the non-randomized design and potential confounding from instructor enthusiasm and motivational differences. Technology acceptance was moderate to strong ($R^2 = .59$ for behavioral intention). Structural equation modeling reveals that perceived usefulness ($\beta = .52$) and instructor guidance ($\beta = .23$) were associated with learning outcomes. Qualitative findings highlight GenAI's potential role as a cognitive scaffolding tool while identifying challenges in preventing over-reliance. This study offers, theoretically, a preliminary conceptual integration of existing frameworks tailored to GenAI-enhanced legal education, and, practically, evidence-based strategies for legal educators to redesign pedagogy and assessment in the AI era.

Keywords Generative artificial intelligence · Legal education · Property law · Pedagogical innovation · Technology acceptance model · TPACK framework

✉ Haifang Liu
ogyz3499@outlook.com

[1] School of Humanities and Law, Hefei University of Technology, Hefei 230000, China

Introduction

Generative artificial intelligence (GenAI), particularly large language models, has triggered substantial changes in higher education since late 2022 (Chan & Colloton, 2024). These tools can generate human-like text, facilitate dialogue, provide immediate feedback, and support problem-solving. Such capabilities offer promising opportunities for teaching and learning, but they also raise concerns about pedagogical adaptability, academic integrity, and evolving instructor roles (Grassini, 2023; Lee & Moore, 2024; Rahimi, 2025). Legal education faces a particularly significant inflection point. Traditional methods such as Socratic dialogue and case-based instruction remain foundational for developing legal reasoning (Shaw, 2012). At the same time, the legal profession is rapidly adopting AI tools, creating new demands for pedagogical innovation (Andreeva & Savova, 2024; Demchenko et al., 2021). Legal reasoning is widely recognized as the core competency of legal education. It encompasses issue identification, rule application, analogical reasoning, and persuasive argumentation (Mertz, 2007; Neumann et al., 2021). Yet empirical evidence on how emerging technologies can enhance the development of these skills remains limited, especially in civil law systems outside the Anglo-American tradition (Stuckey, 2009; Sullivan et al., 2007).

Research in higher education has linked GenAI tools to improvements in writing quality, conceptual understanding, and self-regulated learning when used within structured pedagogical frameworks (Borge et al., 2024; Lee & Moore, 2024). These benefits, however, are highly context-dependent. They cannot be generalized across disciplines without careful attention to the cognitive demands and professional standards of each field (Bobula, 2024; Grassini, 2023). Legal education presents distinctive challenges. Students must learn to evaluate AI-generated legal analysis for doctrinal accuracy, identify potential hallucinations in case citations, and recognize the jurisdictional limitations of AI reasoning (Chee et al., 2025; Garingan & Pickard, 2021). These competencies go beyond general AI literacy. They require discipline-specific critical evaluation skills. In addition, traditional written assessments are vulnerable to AI substitution. Scholars have called for performance-based formats that require real-time reasoning and oral argumentation (Dawson, 2020; Khlaif et al., 2025). This call aligns with longstanding arguments for assessment reform in legal education (Stuckey, 2009).

Despite growing scholarly attention, empirical research on GenAI integration in legal education remains sparse. The existing evidence base is concentrated in Anglo-American common law contexts, with limited attention to civil law systems, non-Western institutional cultures, or large-enrollment teaching models (Andreeva & Savova, 2024; Chimbwanda, 2022). Chinese legal education operates within a civil law tradition and an examination-oriented assessment culture. It presents a contextually distinct setting where GenAI integration may function differently from what has been observed in Western contexts. At the theoretical level, existing frameworks have not kept pace with practice. Conversational AI has distinctive features, including its capacity for Socratic dialogue simulation, iterative feedback, and multi-perspective analysis. These features require discipline-specific theorization that general-purpose educational technology frameworks do not adequately provide (Granić, 2022; Han & Sa, 2022).

The present study addresses these gaps by examining GenAI integration in property law education at a Chinese university. The study is grounded in an integrated framework that combines constructivist learning theory, the Technology Acceptance Model, and the Technological Pedagogical Content Knowledge framework. It investigates three research

questions. First, to what extent is GenAI-enhanced instruction associated with improvements in students' legal reasoning abilities? Second, what factors influence students' adoption intentions toward GenAI technology in legal learning environments? Third, what pedagogical adaptation processes do educators undertake when implementing GenAI integration? By combining quantitative outcome measures with qualitative accounts of implementation processes, this study contributes a preliminary conceptual framework for GenAI-enhanced legal pedagogy and evidence-based strategies for integrating AI tools while maintaining instructional rigor and academic integrity.

Theoretical framework

Constructivist learning theory in legal education

Constructivist theory believes that learning is not a passive process in which learners receive information from teachers but an active process where learners construct their own knowledge through social interactions and experiences (Mishra, 2023). This theory is highly relevant in legal education in which the process of developing legal reasoning needs learners to be intensely engaged with intricate conceptual frameworks, apply abstract principles to practical problems, and build effective arguments on legal concepts and precedents. Vygotsky's (1978) Zone of Proximal Development (ZPD) provides a theoretical basis for GenAI integration in legal education: learners achieve optimal cognitive development when supported through challenging tasks by a more capable agent, a role that GenAI may partially fulfill (Do et al., 2023). Regarding property law education in relation to constructivism theory, GenAI may serve as a mediating platform through which learners engage in interactive case analysis, examine their conceptual assumptions through dialogue, and develop metacognitive awareness of their own reasoning processes (Almulla, 2023).

Technology acceptance model (TAM)

The Technology Acceptance Model serves as a theoretical underpinning in understanding technology adoption (Alshammari & Rosli, 2020). The central components of a TAM—Perceived Usefulness (PU) and Perceived Ease of Use (PEOU)—focus on individual user perceptions of influence in regard to adopting and using technology. In an academic context, an amended version of TAM containing additional dimensions such as subjective norms, facilitating conditions, and self-efficacy acknowledges the complexity of social factors and individual characteristics of an institution in regards to technology adoptions (Teo, 2009). The application of a Technology Acceptance Model to the adoption of GenAI in legal studies would need to consider factors related to discipline-specific perceptions of a convergence of AI capability and legal thought requirements, academic integrity in terms of an over-reliance on AI assistance, and the credibility of legal analysis offered through AI. Acceptance of GenAI technology by students in property law studies may depend upon its ease of use, ability to assist in academic assignments, and an individual belief in its ability to enhance understanding of legal concepts and skills pertinent to a legal career. The amended Technology Acceptance Model used in this investigation examines individual attitude and usage intentions through a mediation role of attitude and behavior

intentions and considering a role of individual teaching and technology affinity as external moderators.

TPACK framework

The Technological Pedagogical Content Knowledge (TPACK) framework identifies the integrated knowledge necessary for effective technology use in teaching (Paskins, 2023). The framework comprises three core knowledge domains: Technological Knowledge (TK), referring to digital tool competence; Pedagogical Knowledge (PK), encompassing teaching practices and learning processes; and Content Knowledge (CK), referring to subject matter expertise. These three domains intersect to produce four additional knowledge levels (Leijen et al., 2022).

In the context of GenAI integration in property law education, each intersection level carries specific meaning. Technological Pedagogical Knowledge (TPK) refers to understanding how to use GenAI tools to support specific teaching strategies, such as using conversational AI to simulate Socratic questioning or to facilitate guided case analysis. Technological Content Knowledge (TCK) refers to understanding how GenAI platforms represent and process legal concepts, including awareness of how AI-generated property law analysis may differ from doctrinally accurate reasoning. Pedagogical Content Knowledge (PCK), widely recognized in legal education scholarship, refers to instructors' knowledge of how to teach legal reasoning skills effectively through appropriate instructional strategies.

The fully integrated TPACK level requires instructors to simultaneously draw on all three domains. In the present study, this means understanding the capabilities and limitations of iFlytek Spark, knowing how to design AI-mediated learning activities that develop legal reasoning, and possessing sufficient property law expertise to evaluate AI-generated outputs for accuracy and disciplinary appropriateness (Kholid et al., 2023). This integrated knowledge cannot be assumed to develop automatically from expertise in any single domain. It requires deliberate professional development that addresses the intersection of technology, pedagogy, and content knowledge (Li, 2022).

Integrated AI-enhanced legal pedagogy framework

The AI-Enhanced Legal Pedagogy Framework proposed in this study does not treat constructivism, TAM, and TPACK as parallel or independent theoretical layers. Rather, each framework performs a distinct function within the integrated model, and the three frameworks constrain and inform one another at specific theoretical junctures. As depicted in Fig. 1, the framework represents these inter-framework relationships through directional linkages connecting the Technological, Pedagogical, Content, and Acceptance dimensions.

Each framework governs a specific aspect of the study. Constructivist learning theory governed the intervention design. Vygotsky's (1978) Zone of Proximal Development provided the rationale for the staged, progressively complex structure of the 16-week GenAI intervention detailed in Table 1. GenAI roles shifted from information provider to Socratic dialogue partner and then to collaborative argumentation tool, reflecting the trajectory from other-regulated to self-regulated performance. TAM governed the measurement framework. Its core constructs of Perceived Usefulness and Perceived Ease of Use provided the theoretical basis for the survey instrument and the structural equation modeling analysis. TPACK governed instructor preparation and implementation consistency, ensuring that

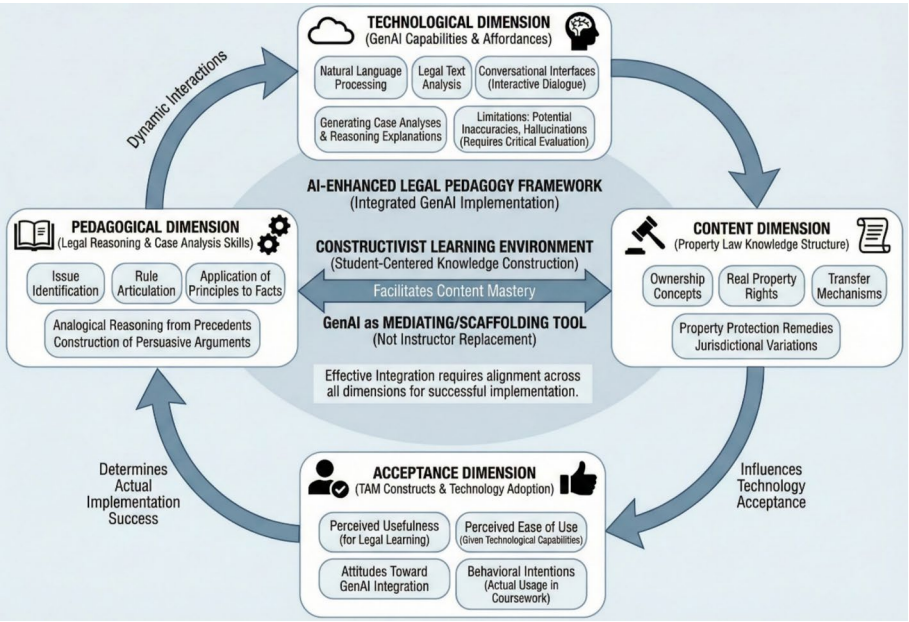

Fig. 1 Conceptual framework of AI-enhanced legal pedagogy

instructors possessed the integrated technological, pedagogical, and content knowledge necessary for effective GenAI integration in property law.

The three frameworks also produce theoretical constraints on one another. TPACK's Technological Content Knowledge dimension directly influenced the TAM construct of Perceived Usefulness. Students' perceptions of whether iFlytek Spark was useful depended in part on their instructors' ability to demonstrate appropriate applications. This means that technology acceptance in professional education cannot be modeled as a purely individual psychological process. The constructivist ZPD concept also shaped how TPACK's Technological Pedagogical Knowledge should be applied. Optimal scaffolding requires calibration to learners' developmental levels, so instructors needed adaptive judgment about when to allow GenAI support and when to redirect students toward independent analysis. In turn, TAM's Behavioral Intention construct fed back into the constructivist framework. Voluntary and intentional GenAI engagement was a prerequisite for the active knowledge construction that constructivism predicts will produce durable learning gains. This integrated framework is offered as a preliminary conceptual contribution subject to refinement through replication across diverse legal education settings.

Research hypotheses

Based on the integrated framework, eight hypotheses were developed to examine relationships among technology acceptance, instructional variables, and learning outcomes.

H1, grounded in constructivist scaffolding principles, predicts that GenAI-enhanced instruction is positively associated with legal reasoning ability, such that students in the

Table 1 Intervention design matrix: GenAI integration in property law course

Week	Learning objectives	Activities	GenAI role
1–4	Understand GenAI capabilities; identify property law issues; apply basic ownership principles	Introduction sessions; simple case analysis with iFlytek Spark; discussion of AI limitations; academic integrity training	Information provider; reasoning demonstrator; interactive tutor for basic concepts
5–8	Analyze complex property rights conflicts; evaluate competing claims; apply temporal and relational considerations	Multi-party property disputes; Adverse possession scenarios; Co-ownership problems; hypothesis testing with AI	Socratic dialogue partner; alternative perspective generator; analytical reasoning scaffold
9–12	Construct legal arguments; develop persuasive writing; conduct legal research; evaluate argument quality	Drafting legal opinions; Peer review with AI feedback; counterargument development; citation assistance	Writing assistant; feedback provider; research aid; argument critic
13–16	Synthesize property law knowledge; prepare oral advocacy; evaluate case theories; perform under pressure	Moot court case preparation; examination question development; mock oral arguments; final assessment	Strategic planning assistant; opposing counsel simulator; performance preparation tool

GenAI condition will demonstrate greater reasoning gains than those receiving traditional instruction.

H2, extending TAM to the legal education context, predicts that perceived usefulness positively influences attitude toward GenAI use, such that students who perceive GenAI as valuable for legal learning are more likely to develop favorable orientations toward its integration.

H3, following established TAM principles, predicts that perceived ease of use positively influences perceived usefulness, such that students who find GenAI interfaces accessible are more likely to recognize their value for legal learning (Al-Adwan et al., 2023).

H4, also drawing on TAM, predicts that attitude toward GenAI use positively influences behavioral intention, such that students with more favorable attitudes are more likely to intend continued GenAI engagement.

H5 predicts that instructor guidance positively moderates the relationship between GenAI use and learning outcomes, such that structured instructor support amplifies the learning benefits of GenAI integration (Jiménez Sierra et al., 2023).

H6 predicts that prior technology experience positively influences perceived ease of use, such that students with greater technology familiarity encounter fewer difficulties when engaging with GenAI tools (Lin & Yu, 2023).

H7 predicts that instructor endorsement positively influences perceived usefulness, such that instructor advocacy for GenAI integration enhances students' perceptions of its value for legal learning (Mulyani et al., 2025).

H8 predicts that behavioral intention positively predicts actual GenAI use, such that students who express stronger intentions to use GenAI are more likely to engage with it in practice.

Methodology

Research design

This study employed a sequential explanatory mixed-methods design to examine GenAI integration in legal education (Love et al., 2022). The quantitative component used a quasi-experimental design with non-equivalent control groups. Learning outcomes and technology acceptance were compared between students receiving GenAI-integrated instruction and those receiving traditional instruction in property law courses at a Chinese research university. Random assignment was not feasible due to intact class section structures, necessitating the quasi-experimental approach. The qualitative component followed the quantitative intervention and used interviews, reflection journals, and classroom observations to explain how and why the observed effects occurred. This integrated approach provided both performance evidence and process-level accounts of GenAI integration.

Participants

The study involved 180 third-year law students enrolled in property law courses during the fall semester of 2023. Students were assigned to the experimental group (n = 90) or the control group (n = 90) based on natural class sections to reduce contamination effects and ensure educational consistency. The experimental group included 52 female students (57.8%) and 38 male students (42.2%), with a mean age of 21.2 years (SD = 0.8). The

control group included 49 female students (54.4%) and 41 male students (45.6%), with a mean age of 21.4 years (SD = 0.9). Baseline equivalence between groups was verified through pre-test legal reasoning scores and prior legal GPA, as reported in "Descriptive statistics and baseline equivalence" section.

For the qualitative component, 30 students were purposively sampled from the experimental group to represent varied levels of technology acceptance, academic performance, and engagement. Three law lecturers who delivered the experimental group instruction participated in in-depth interviews about their implementation experiences and pedagogical adaptations. The control group was taught by three different instructors with comparable qualifications, all holding doctoral degrees in law and 8–12 years of teaching experience.

Ethics approval was granted by the Institutional Review Board of Hefei University of Technology. Ethical safeguards included participant anonymization, protection from academic disadvantage, and provision of post-study GenAI access for all students. Because the intervention required students to submit legal analyses to iFlytek Spark, a commercial platform, additional data governance measures were implemented. All participants provided informed consent addressing data privacy. The research team anonymized all interaction data and stored it on password-protected institutional servers. Students were instructed to use only hypothetical case scenarios and to avoid including personally identifiable information in their submissions. The platform's data processing practices were reviewed against China's Personal Information Protection Law. However, complete elimination of platform-level data retention risk cannot be guaranteed, and this limitation is discussed in "Study limitations" section.

GenAI intervention design

The intervention systematically integrated iFlytek Spark, a commercially available GenAI platform developed by iFlytek Co., Ltd. with Chinese-language processing capabilities, into property law teaching over a 16-week semester comprising 48 contact hours. The platform was selected based on language appropriateness for Chinese legal education, ease of use, and compatibility with national data governance norms.

The intervention followed a model of progressively increasing complexity. As shown in Table 1, each stage was designed with specific learning objectives, activity types, and GenAI role expectations. Figure 2 illustrates the three-phase instructional workflow implemented in each class session during the intervention. GenAI roles shifted from information provider and reasoning demonstrator in Weeks 1–4, to Socratic dialogue partner in Weeks 5–8, to writing assistant and feedback provider in Weeks 9–12, and finally to strategic planning assistant and opposing counsel simulator in Weeks 13–16. The control group received the same learning materials through conventional instruction, including lectures, faculty-led case discussions, and collaborative tasks without GenAI integration.

Several measures minimized potential AI contamination in the control group. Control group instructors explicitly requested that students refrain from using GenAI tools for course assignments. Course assessments emphasized in-class performance, including oral defenses, timed analyses, and moot court performances, rather than solely take-home assignments. Post-study surveys indicated that 89% of control group students reported no or minimal GenAI use during the semester, though this self-reported measure represents a limitation discussed in "Study limitations" section. Two additional structural mechanisms supplemented self-report data. Classroom observations across eight sessions recorded no evidence of AI tool engagement during class time. The assessment design itself functioned

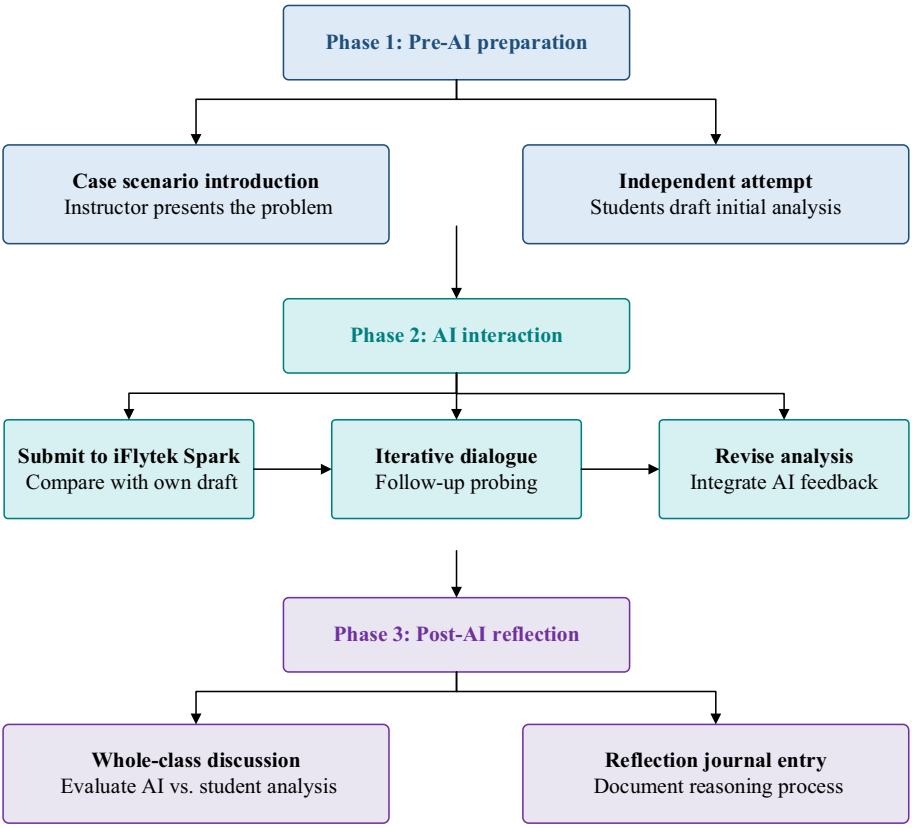

Fig. 2 Three-phase instructional workflow per session

as a safeguard, as performance-based formats required real-time legal reasoning that cannot be meaningfully substituted by prior AI-assisted preparation alone. These convergent lines of evidence collectively support the conclusion that differential AI exposure between groups was sufficiently controlled.

Data collection instruments

Quantitative data were collected using three instruments. The Legal Reasoning Ability Test (LRAT) was a 30-item pre-post test evaluating five domains of legal reasoning in property law: issue identification, rule articulation, application of principles to facts, analogical reasoning from precedent, and persuasive argument construction. Each item presented a written property law scenario requiring written analysis. Responses were scored on a 0 to 100 scale using a validated assessment rubric.

The LRAT was developed through a multi-stage validation process. An initial pool of 48 scenario-based items was constructed in consultation with five external property law faculty members. Items were classified according to a two-dimensional taxonomy distinguishing doctrinal recall from analytic reasoning. Following expert review, 42 items were retained. A pilot administration with 45 non-participating third-year law students confirmed

satisfactory item difficulty indices (.42 to .71) and item-total correlations exceeding .30. The final 30-item battery was then selected. Parallel pre-test and post-test forms were constructed using a matched-pairs design. Each item pair was matched on legal domain, cognitive demand, and difficulty index (difference ≤ .05). Five anchor items appeared on both forms to enable statistical equating. Confirmatory factor analysis on pilot data supported a unidimensional structure (CFI = .93, RMSEA = .06) and measurement equivalence across forms (ΔCFI < .01). Two trained researchers scored all responses blind to group assignment. Inter-rater reliability on 30% of randomly selected responses yielded an intraclass correlation of .87. Disagreements exceeding five points were resolved by a third rater. A complete sample item with scoring rubric is provided in Appendix A.

The Technology Acceptance Questionnaire (TAQ) was a 35-item self-report measure using a seven-point Likert scale. It assessed six constructs from an extended TAM: Perceived Usefulness, Perceived Ease of Use, Attitude Toward Use, Behavioral Intention, Instructor Guidance, and Prior Experience. An additional subscale addressed academic integrity concerns. Core PU and PEOU items were adapted from Davis (1989). ATU and BI items were adapted from Teo (2009). IG and PE items drew from Al-Adwan et al. (2023). Academic integrity items were developed by the research team and subjected to expert review. All items underwent a structured localization process involving forward translation into Mandarin Chinese, review by two property law faculty members and one educational technology specialist, and independent back-translation. Generic usefulness statements were reworded to reflect legal professional competency contexts. A pilot with 32 non-participating students confirmed satisfactory internal consistency across all subscales (α = .80 to .88). Course grades and assignment scores were also collected as learning outcome measures.

Qualitative data were collected through four methods. Semi-structured interviews were conducted with 30 students and three instructors. Each interview lasted 45 to 60 min and covered experiences with GenAI integration, perceived benefits and challenges, and suggestions for improvement. Students completed bi-weekly reflection journals eight times during the semester, recording specific instances of GenAI use, learning outcomes, and evolving attitudes. Classroom observations followed a structured 45-min protocol across eight sessions, documenting instructional strategies, student engagement, and technology use. GenAI usage logs from the iFlytek Spark analytics dashboard recorded interaction frequency, duration, and query types.

Data analysis procedures

Quantitative data analysis employed SPSS 26.0 for descriptive statistics and independent samples t-tests comparing experimental and control groups on legal reasoning gains, with effect sizes calculated using Cohen's d formula

$$d = \frac{M_1 - M_2}{SD_{pooled}} \tag{1}$$

where $SD_{pooled} = \sqrt{\frac{(n_1-1)SD_1^2+(n_2-1)SD_2^2}{n_1+n_2-2}}$. Analysis of covariance (ANCOVA) controlled for pre-test scores when examining post-test differences. Structural equation modeling using Smart PLS 4.0 tested the extended TAM framework through two-stage analysis (Guo, 2024). The SEM analysis was conducted on the full study sample of 180 participants rather than the experimental group alone, as the TAM-based structural model—encompassing

technology acceptance perceptions, behavioral intentions, and actual use behaviors—was assessed across all participants who completed the Technology Acceptance Questionnaire regardless of group assignment. This decision was made on the grounds that technology acceptance constructs represent individual psychological responses applicable to all students who had exposure to the research context, and that restricting SEM analysis to the experimental group (n=90) would have yielded a sample-to-parameter ratio insufficient to support stable estimation of a seven-construct latent variable model. With 180 participants and a model comprising seven latent constructs and 40 observed indicators, the present analysis achieves a cases-to-indicators ratio of 4.5:1, which falls within the range considered acceptable for PLS-SEM estimation (Hair et al., 2019), though it should be noted that this ratio approaches the lower boundary of recommended thresholds for CB-SEM approaches. The use of PLS-SEM via SmartPLS 4.0 was therefore particularly appropriate given its relative robustness with smaller sample sizes compared to covariance-based SEM alternatives (Ringle et al., 2015). Measurement model evaluation examined indicator loadings (threshold >.70), internal consistency via Cronbach's alpha and composite reliability $CR = \frac{(\sum \lambda_i)^2}{(\sum \lambda_i)^2 + \sum var(\epsilon_i)}$ (threshold >.70), convergent validity through average variance extracted $AVE = \frac{\sum \lambda_i^2}{n}$ (threshold >.50), and discriminant validity using Fornell-Larcker criterion. Structural model assessment employed bootstrapping with 5000 resamples to evaluate path coefficients and R^2 values. Qualitative data analysis utilized NVivo 14 for systematic thematic analysis following open, axial, and selective coding procedures (Allsop et al., 2022). Two trained coders independently coded 30% of data, achieving Cohen's kappa of .83. Intercoder disagreements were resolved through a structured two-stage reconciliation process. In the first stage, coders independently reviewed all discrepant items and provided written rationales for their coding decisions; items where disagreement reflected differing interpretations of theme boundaries were resolved through direct discussion and consensus, with the agreed resolution recorded in a shared codebook annotation. In the second stage, items where consensus could not be reached through discussion alone—representing approximately 4% of all coded segments—were, submitted to a third qualified researcher with expertise in educational technology for adjudication, whose coding decision was accepted as final. This process ensured that the final coding scheme reflected systematic analytical judgment rather than the perspective of any single coder. Data integration followed sequential explanatory design logic (Draucker et al., 2020), using joint displays to systematically compare quantitative patterns with qualitative themes, identifying convergence and divergence areas.

Results

Quantitative findings

Descriptive statistics and baseline equivalence

To afford the preliminary analysis, participant characteristics were investigated to ensure that the experimental group and control group were equivalent in every manner to ensure that the main analysis is valid. Based on the findings in Table 2, the experimental group consisted of 90 students where 52 were female (57.8%) and 38 were male (42.2%), ranging in age 21.2 years old (M = 21.2, SD = 0.8), while the control group consisted of 90 students

Table 2 Participant demographics and baseline equivalence

Variable	Experimental group (n = 90)	Control group (n = 90)	Statistical test	p-value
Gender (female/male)	52/38 (57.8%/42.2%)	49/41 (54.4%/45.6%)	$\chi^2 = 0.21$	.647
Age M(SD)	21.2(0.8)	21.4(0.9)	t = 1.52	.131
Pre-test legal reasoning M(SD)	65.4(8.2)	64.8(8.5)	t = 0.48	.632
Prior legal GPA M(SD)	3.42(0.51)	3.38(0.49)	t = 0.53	.596
Technology experience M(SD)	3.6(0.9)	3.5(0.8)	t = 0.79	.432

where 49 were female (54.4%) and 41 were male (45.6%), ranging in age 21.4 years old (M = 21.4, SD = 0.9). Chi-square analysis demonstrated that the proportion of gender in both groups is not significantly distinct ($\chi^2 = .21$, p = .647), while the independent samples T-test demonstrated that the differing ages between the two groups were not significantly distinct (t(178) = 1.52, p = .131). Most importantly, to ensure that the experimental group would not be placed at a disadvantage in comparison to the control group in relation to legal reasoning ability tests, preliminary analysis demonstrated that legal reasoning ability scores were equal between the experimental group M = 65.4 (SD = 8.2) and the control group M = 64.8 (SD = 8.5) with no significant differences (t(178) = 0.48, p = .632). Similarly, GPA in legal studies per se demonstrated equality in the experimental group M = 3.42 (SD = 0.51) in comparison to the control group M = 3.38 (SD = 0.49), where T = 0.53, p = .596. Self-reported levels of previous technology familiarity using a scale of 5 levels demonstrated equivalence in technological familiarity between the experimental group M = 3.6 (SD = 0.9) in comparison to the control group M = 3.5 (SD = 0.8), T = 0.79, p = .432.

Legal reasoning ability development

Results from analysis of legal reasoning ability showed that there were large and statistically significant differences in learning gains between the experimental group and the control group. As shown in Table 3, the experimental group showed significant growth from pre-test (M = 65.4, SD = 8.2) to post-test (M = 78.6, SD = 7.5), which translated to a mean score gain of 13.2. However, the control group showed less significant growth from pre-test (M = 64.8, SD = 8.5) to post-test (M = 69.2, SD = 8.1), which was a mean score gain of 4.4. The independent samples t-test performed to compare post-test scores between the two groups showed that the differences were highly significant, t(178) = 7.82, p < .001, and the

Table 3 Pre-test and post-test legal reasoning scores comparison and between-group effect sizes

Group	Pre-test M(SD)	Post-test M(SD)	Mean Gain	t-value	p-value	Cohen's d
Experimental (n = 90)	65.4(8.2)	78.6(7.5)	13.2	11.34*	< .001	1.68
Control (n = 90)	64.8(8.5)	69.2(8.1)	4.4	3.87*	< .001	0.52
Between-group comparison	t = 0.48	t = 7.82*	–	–	< .001	1.17

*p < .001. Cohen's d values represent within-group effect sizes for pre-post changes (rows 1–2) and between-group effect size for post-test comparison (row 3)

observed effect size (Cohen's d = 1.17) is notably large by conventional criteria (Abrami et al. 1988); however, this magnitude warrants cautious interpretation in the context of a non-randomized quasi-experimental design, as effect sizes of this scale in non-randomized educational intervention studies are susceptible to inflation from confounding sources including instructor enthusiasm, differential student motivation arising from novelty of the experimental condition, and uncontrolled contextual variables that randomization would otherwise distribute equally across groups. ANCOVA controlling for pre-test scores confirmed this pattern, $F(1,177) = 61.24$, $p < .001$, partial $\eta^2 = .26$, indicating that the association between GenAI-enhanced instruction and learning gains remained substantial after accounting for pre-test differences, though ANCOVA controls for initial ability levels only and cannot account for motivational, attitudinal, or instructor-related confounds that may have contributed to the observed effect magnitude. As shown by the results, the experimental group showed particularly significant gains in applying principles to facts (18.5% growth) and developing persuasive arguments (16.2% growth), compared to less significant growth, respectively, in the control group (6.8% and 5.4% growth). As shown by Fig. 3, the experimental group generated faster growth compared to the control group, as demonstrated by the statistically significant group-by-time interaction effect, $F(1,178) = 52.73$, $p < .001$. The results offer preliminary evidence consistent with hypothesis one, which predicted that GenAI-enhanced instruction would be associated with greater learning gains regarding legal reasoning abilities in property law education.

Technology acceptance model results

Structural equation modeling analysis provided preliminary support for the extended TAM framework in explaining GenAI adoption among law students. Assessment of the measurement model revealed excellent psychometric properties across all constructs, as detailed in Table 4. Indicator loadings ranged from .72 to .89, substantially exceeding the recommended threshold of .70, suggesting adequate relationships between constructs and their respective indicators. Internal consistency reliability was demonstrated through Cronbach's alpha coefficients ranging from .82 to .91 and composite reliability values between .85

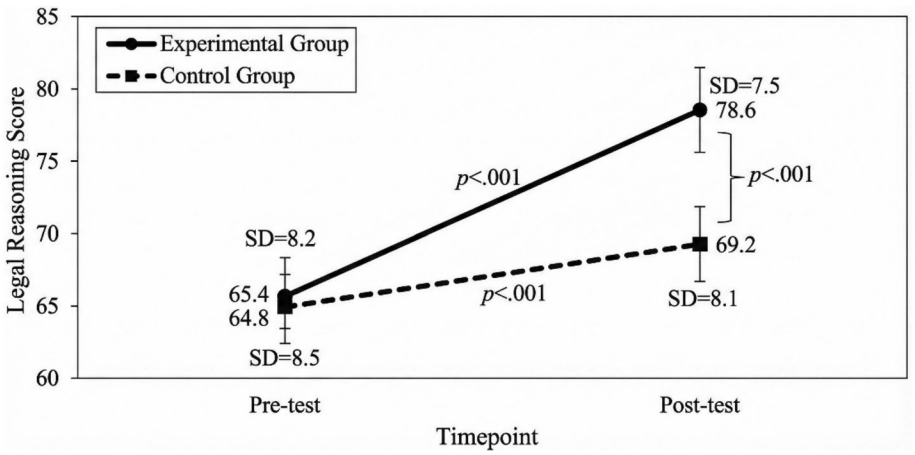

Fig. 3 Legal reasoning score changes over time

Table 4 Measurement model assessment

Construct	Items	Loadings range	Cronbach's α	CR	AVE
Perceived usefulness (PU)	8	.74–.87	.89	.91	.68
Perceived ease of use (PEOU)	7	.72–.85	.86	.88	.64
Attitude toward use (ATU)	6	.78–.89	.91	.93	.72
Behavioral intention (BI)	5	.81–.88	.88	.90	.70
Instructor guidance (IG)	4	.76–.84	.82	.85	.66
Prior experience (PE)	5	.73–.86	.84	.87	.67
Legal reasoning ability (LRA)	5	.75–.87	.87	.89	.78

CR Composite reliability, *AVE* average variance extracted

and .93, all surpassing the .70 criterion. Convergent validity received support from average variance extracted (AVE) values spanning .64 to .78, with all constructs exceeding the .50 threshold. Discriminant validity was established through the Fornell-Larcker criterion, with the square root of each construct's AVE exceeding its correlations with other constructs, while HTMT ratios remained below .85.

Structural model evaluation, depicted in Fig. 4 and detailed in Table 5, revealed differential levels of support for the eight hypothesized relationships. Core theoretical pathways demonstrated highly significant effects. GenAI integration was significantly associated with legal reasoning ability (H1: $\beta = .43$, $t = 6.21$, $p < .001$), providing preliminary evidence consistent with the predicted positive relationship between GenAI-enhanced instruction and reasoning development. Within the TAM framework, perceived usefulness strongly predicted attitude toward use (H2: $\beta = .52$, $t = 7.84$, $p < .001$), while attitude significantly influenced behavioral intention (H4: $\beta = .61$, $t = 9.15$, $p < .001$). Behavioral intention strongly predicted actual usage (H8: $\beta = .58$, $t = 8.32$, $p < .001$), consistent with the theoretical chain from perceptions to behaviors.

Contextual and moderating variables showed significant but comparatively weaker effects. Perceived ease of use positively influenced perceived usefulness (H3: $\beta = .32$,

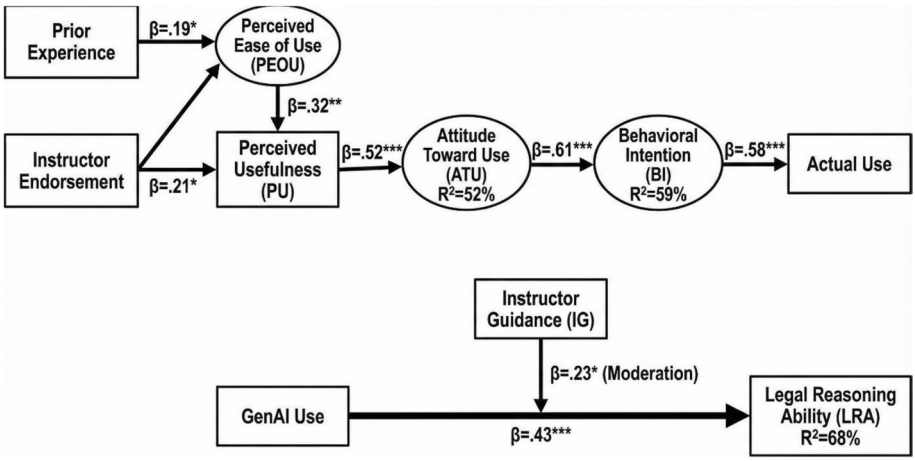

Fig. 4 Structural equation model

Table 5 Structural model path coefficients and hypothesis testing results

Hypothesis	Path	β	t-value	p-value	95% CI	Decision
H1	GenAI use → legal reasoning	.43	6.21	< .001	[.29, .56]	Supported
H2	PU → ATU	.52	7.84	< .001	[.39, .64]	Supported
H3	PEOU → PU	.32	2.91	.004	[.11, .52]	Supported
H4	ATU → BI	.61	9.15	< .001	[.48, .73]	Supported
H5	IG × GenAI → learning	.23	2.38	.018	[.04, .42]	Supported
H6	Prior experience → PEOU	.19	2.05	.042	[.01, .37]	Supported
H7	Instructor endorsement → PU	.21	2.17	.031	[.02, .39]	Supported
H8	BI → actual use	.58	8.32	< .001	[.44, .71]	Supported

PU Perceived usefulness, *PEOU* perceived ease of use, *ATU* attitude toward use, *BI* behavioral intention, *IG* instructor guidance

t = 2.91, p = .004), instructor guidance moderated learning outcomes (H5: β = .23, t = 2.38, p = .018), prior technology experience affected ease of use perceptions (H6: β = .19, t = 2.05, p = .042), and instructor endorsement contributed to perceived usefulness (H7: β = .21, t = 2.17, p = .031). This pattern of differential significance suggests that student-centered perceptual factors exert stronger influence than external contextual variables in GenAI acceptance. The model explained substantial variance in endogenous constructs: 52% in attitude toward use, 59% in behavioral intention, and 68% in legal reasoning ability. Model fit indices demonstrated excellent alignment with observed data (SRMR = .062, NFI = .91, CFI = .94), confirming the appropriateness of the extended TAM framework for understanding GenAI adoption in legal education contexts.

Learning outcomes analysis

Analysis of learning outcomes beyond legal reasoning test scores revealed a persistent benefit of the experimental group across a variety of assessment metrics. For instance, as depicted in Table 6, the experimental group earned a significantly higher average grade on course assignments with a focus on case analysis (experimental group, M = 82.4, SD = 6.8, vs. Control Group, M = 76.1, SD = 7.4), t(178) = 5.93, p < .001, Cohen's d = 0.88. Performance on legal writing assignments revealed a significantly higher average grade in the experimental group (M = 79.6, SD = 7.2) as opposed to the control group (M = 73.8, SD = 7.9), t(178) = 5.14, p < .001, Cohen's d = 0.76. Performance on the final examination, composed of traditional doctrinal as well as practical reasoning questions,

Table 6 Course performance comparison between groups

Assessment component	Experimental M(SD)	Control M(SD)	t-value	p-value	Cohen's d
Case analysis assignments	82.4(6.8)	76.1(7.4)	5.93	< .001	0.88
Legal writing assignments	79.6(7.2)	73.8(7.9)	5.14	< .001	0.76
Moot court performance	80.8(8.1)	75.3(8.6)	4.37	< .001	0.65
Final examination	81.2(8.4)	74.5(8.9)	5.19	< .001	0.77
Overall course grade	81.0(7.6)	74.9(8.2)	5.23	< .001	0.78

was also significantly greater in the experimental group (M = 81.2, SD = 8.4) as opposed to the control group (M = 74.5, SD = 8.9), t(178) = 5.19, p < .001, Cohen's d = 0.77. Analysis of course grade distributions (depicted in Fig. 5) also revealed a statistically significant tendency of 67% of experimental subjects to earn a course grade of B+ and higher as opposed to 42% of controls (χ^2 = 11.34, p = .001). Further examination of findings revealed the experimental group to excel in terms of the efficacy of legal reasoning (18.5% greater than controls), completeness of analysis (15.2% greater), and persuasiveness of argumentation (14.7% greater) while a lesser margin existed with regards to knowledge recall (8.3% greater).

Qualitative findings

Student perspectives

Thematic analysis of interviews and reflection journals revealed three broad themes regarding GenAI-assisted legal education. The dominant theme, enhanced legal understanding through AI-driven education, comprised 42% of all coded citations. Students emphasized how iFlytek Spark's immediate feedback enabled iterative learning. They could test hypotheses, receive instant responses, and refine their analyses through multiple rounds without waiting for instructor input. Students also found GenAI useful for generating multiple legal perspectives on complex cases. This exposure helped them recognize that legal reasoning involves evaluating competing interpretations rather than arriving at a single correct answer. As one student explained, "When using iFlytek Spark for property law, I can test my legal reasoning and receive answers within an instant. If my reasoning was faulty, I simply revisit it, make new determinations, and test again, all without needing days for professors' input." Students who engaged most actively in this iterative process tended to

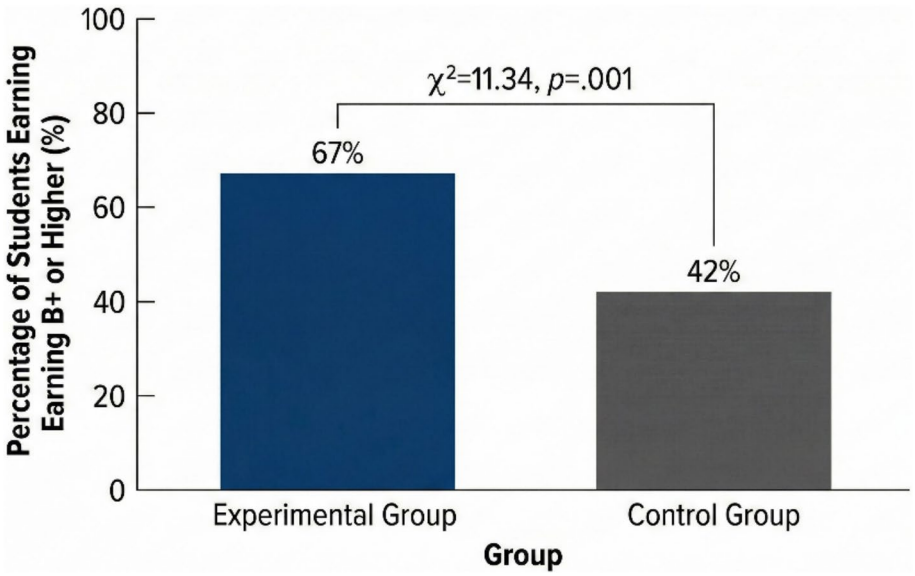

Fig. 5 Distribution of final course grades

perform better on post-test evaluations. Figure 6 presents authentic iFlytek Spark interface screenshots showing a representative revision trajectory, illustrating how one student refined a legal analysis through iterative engagement with platform feedback across three successive exchanges.

The second theme, confidence building and self-regulated learning, emerged from 31% of coded citations. Students described GenAI as a low-stakes practice environment where they could develop their skills without fear of judgment. One student noted, "I can practice case analysis using iFlytek Spark without fearing embarrassment or disappointing my teacher. Private failure has been crucial to my success in public performance." GenAI also encouraged self-evaluation. Students reported developing criteria for assessing argument quality, identifying gaps in their analyses, and recognizing when further research was needed. These patterns are consistent with Vygotskian scaffolding principles and the progression from external support to autonomous performance.

The third theme, critical evaluation and over-reliance concerns, accounted for 27% of coded citations. Roughly 23% of students cautioned against excessive dependence on GenAI. They recognized that AI-generated outputs require verification against course materials and legal knowledge. However, usage log data also revealed a more subtle risk. A



(a) Initial draft — Before AI interaction

(b) Revised draft — After first AI exchange

(c) Final version — After second AI exchange

Fig. 6 Authentic iFlytek Spark interface showing a representative student revision trajectory through iterative AI feedback engagement. Screenshots captured from the iFlytek Spark mobile application using the same hypothetical scenario described in Appendix A. Interface in Mandarin Chinese. **a** Initial draft before AI interaction; **b** Revised draft after first AI exchange incorporating statutory citations; **c** Final version after second AI exchange demonstrating multi-perspective analysis and acknowledgment of legal weakness. Student data are composite; scenario is hypothetical

subset of students, submitted three or more successive revisions to the same prompt within a single session, with each revision incorporating more of the AI's structural features and terminology. Approximately 18% of reflection journal entries contained statements suggestive of conformity-oriented revision behavior, such as adjusting answers until the platform provided a more positive response. This pattern of algorithmic conformity is discussed in detail in "Discussion" section.

Several negative cases complicated the predominantly positive pattern. First, some students used GenAI in a surface-level or confirmatory manner without engaging with the explanations provided. One student reflected, "I mostly used iFlytek Spark to check whether my answer was roughly correct before submitting." This suggests that the scaffolding affordances of GenAI were not uniformly activated across the experimental group. Second, several students expressed active resistance to GenAI integration, viewing it as a shortcut that could weaken independent legal reasoning. Third, a distinct over-reliance pattern was observed among students who consulted the platform before attempting independent analysis rather than after. This raises questions about whether the intervention successfully positioned GenAI as a scaffold to be progressively withdrawn. These negative cases are incorporated into the integrated analysis in "Integrated analysis" section, and their theoretical implications are addressed in "Discussion" section.

Instructor experiences

Analysis of the interviews with the teachers reveals four interrelated themes that capture the collective experience of the teachers with the application of GenAI-enhanced teaching. It should be noted that instructor interview data were drawn from a purposive sample of three law lecturers involved in delivering the experimental group instruction, representing a small sample that limits the transferability of instructor-level findings. While the depth of qualitative engagement with each instructor—including multiple interview sessions and classroom observation records—partially compensates for the limited sample size, the instructor themes reported below should be interpreted as illustrative of individual practitioner experiences within this specific institutional context rather than as representative of legal educators more broadly. Transformation of the teaching role appears to emerge from the integration of GenAI into the teaching framework. All the teachers state that they have moved from the role of mere information providers to facilitators of the teaching process that uses AI for knowledge development. As one teacher said, "Since routine explanations and feedback can be done by the iFlytek Spark teaching assistant, I am now able to devote my time to high-order tasks such as preparing students to analyze the fallacies involved in AI's reasoning processes, think up policies that align with AI's outputs, and model more advanced legal argumentation that surpasses AI's capability."

Second theme: assessment design. Lecturers are less likely to rely on traditional assessments that are more likely to require AI. They rather choose performance tasks assessing HOTS, something that is hard to achieve with GenAI. Table 7 shows this. One of the lecturers believed that some analyses that can easily be done by AI do not measure ability. They also require an oral argument, immediate problem resolution, or combining various legal fields, in some of which AI can prepare but not substitute intelligence.

The third theme, professional development needs, refers to the concerns of teachers regarding the basics, limitations, and proper use of GenAI. They discussed technology ignorance regarding AI platforms, reservations regarding effective applications of teaching, and concerns regarding the control of instruction in the classroom due to AI use by

Table 7 Assessment transformation: traditional versus AI-enhanced approaches

Traditional assessment	AI-enhanced assessment	Rationale
Written case analysis (take-home)	Oral case defense with follow-up questions	Requires real-time reasoning beyond AI capability
Multiple-choice doctrinal questions	Scenario-based problem-solving (timed, in-class)	Emphasizes application under constraints
Individual research papers	Collaborative moot court with process portfolio	Assesses teamwork and metacognitive awareness
Final comprehensive exam	Multi-stage project with peer review and revision	Values iterative development and feedback integration

students. Each emphasized professional development that included technical education and pedagogic innovation, as well as collaboration, for mutual support and strategy on use. A teacher observed, "A new skill set for GenAI use is required. There must be institutional support for us not just to use but to teach."

The fourth theme: ethical and academic integrity issues reflects the struggle teachers undergo to balance GenAI's positive in-class effects and the prospect of cheating. They described establishing usage guidelines, prompting students to record their use and learning contribution of AI, and designing assessment strategies to minimize dependence on AI. However, teachers find it challenging to suppress cheating aided by AI, requiring frequent revisions in integrity guidelines and strategies in assessment. One teacher said, "We need to train students on proper AI usage, not prohibit it outright. The future of the legal practice will depend on working with AI; therefore, students need to learn to use this reality appropriately."

GenAI interaction patterns

There are identifiable usage patterns on the part of the students using the iFlytek Spark system for property law study. Table 8 shows that case analysis help usage dominates with a usage rate of 42%, with students submitting case scenario queries asking for appropriate issues, rules, and framework applications. Socratic dialogue simulation usage follows with 38%, involving elaborate consultations with significant AI-probing queries. Legal writing feedback usage contributes 28%, whereby students submit writings requiring feedback on organization, persuasiveness, and completeness. Research assistance usage follows with 19%, concentrating on identifying principles, precedents, or sources. Concept clarification usage contributes 15%, involving the elucidation of intricate property law concepts or terms. Several usages may also take place simultaneously with the same group of students. For instance, they may employ concept clarification usage initially, then case analysis help usage, finishing with socratic dialogue simulation usage. Efficient patterns, which reflect learning outcomes, incorporate dialogue interactions rather than the passive analysis of AI, suggesting the importance of engagement with GenAI tools as learning partners over just using the tools as resources to answer questions.

Integrated analysis

Systematic integration of findings provides preliminary support for the AI-Enhanced Legal Pedagogy Framework and the nature of those effects. From the quantitative study, perceived usefulness was a strong predictor of technology acceptance ($\beta = .52$). Effectiveness from the qualitative study was linked back to iterative learning (42% of coded statements). The feedback from the GenAI tool allowed for the speedy improvement of work with multiple attempts at analytical reasoning. The moderator role of teaching support ($\beta = .23$) was consistent with themes around the provision of scaffolding instruction on usage and metacognitive strategies that guard against overuse and promote effective AI usage. Student performance was better with high teaching support.

Highly variable variance in legal reasoning ability ($R^2 = .72$) was consistent with the qualitative understanding of the support mechanisms in GenAI, namely, intellectual scaffolding in dialogue, analysis familiarization, reduced test anxiety in the practice phase, and metacognitive enhancement from comparison of one's own analysis and the analysis generated by the AI system (Fig. 7). Qualitative data triangulated the validity of the conceptual

Table 8 GenAI usage type coding framework and frequency distribution

Usage type	Description	Frequency (%)	Effectiveness rating[a]
Case analysis assistance	Requesting issue identification, rule application, analysis structure	42	8.2/10
Socratic dialogue simulation	Interactive questioning to test reasoning and explore alternatives	38	9.1/10
Legal writing feedback	Submitting drafts for critique on organization and persuasiveness	28	8.6/10
Research assistance	Querying relevant legal principles, precedents, sources	19	7.4/10
Concept clarification	Requesting explanations of complex concepts or terminology	15	7.8/10

[a]Effectiveness rating based on correlation with learning outcome gains and student-reported perceived utility

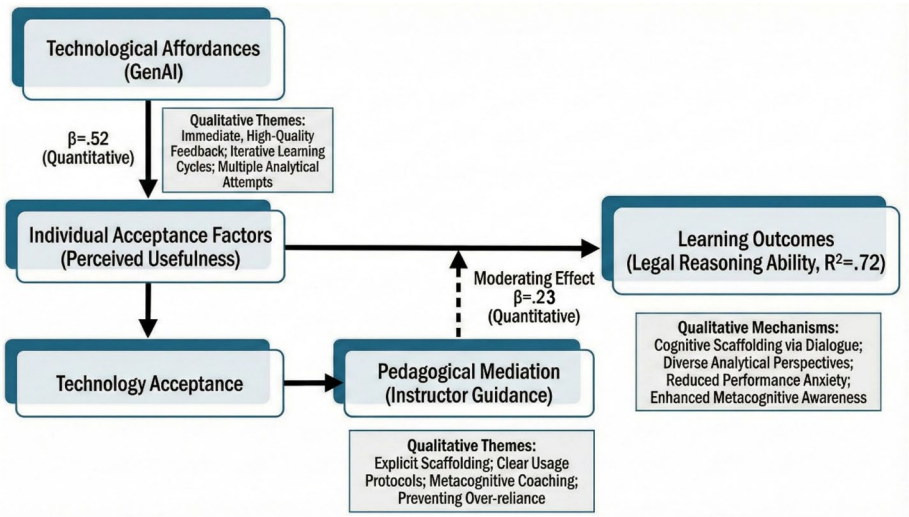

Fig. 7 Integrated findings model

model, except in the perception of academic integrity, where qualitative findings suggested a more complex level of consideration of the concept in the questionnaire, as the data indicated little test anxiety, in contrast to the expectations from qualitative insights.

Discussion

The observed between-group difference in legal reasoning gains (Cohen's d = 1.17) suggests a meaningful association between GenAI-enhanced instruction and improved legal reasoning outcomes. However, this effect size is notably large relative to benchmarks typically reported in non-randomized educational intervention research (Hyde-Vaamonde & Keller, 2025). Three sources of potential inflation cannot be ruled out. Instructor enthusiasm effects may have operated differentially across conditions, as experimental group instructors were directly involved in designing the intervention and may have exhibited greater pedagogical investment independently of the GenAI component. Motivational differences arising from the novelty of the experimental condition may have elevated engagement beyond levels attributable to GenAI scaffolding alone. The non-randomized assignment of intact class sections also means that unobserved between-class differences may have contributed to the performance gap. The effect size should therefore be understood as an upper-bound estimate pending replication with random assignment and active control of instructor-level variables.

Regarding the novelty concern specifically, novelty effects are typically characterized by an initial engagement surge that diminishes within four to six weeks (Cheung & Vogel, 2013; Clark, 1983). The present intervention spanned a full 16-week semester. As illustrated in Fig. 3, the experimental group demonstrated a sustained upward trajectory across the entire period rather than early acceleration followed by plateau or decline. This temporal pattern is inconsistent with a novelty-driven effect. Qualitative data from weeks 13–16 also documented continued substantive engagement with iterative reasoning processes

rather than reduced interest. While these observations do not entirely eliminate the novelty explanation, they suggest that the observed gains reflect pedagogical benefit rather than transient motivation.

The qualitative findings suggest that GenAI may function as a cognitive scaffolding tool consistent with Vygotskian theory (Kim & Adlof, 2024). Students benefited from immediate feedback, exposure to multiple analytical perspectives, and low-stakes practice opportunities. Importantly, the primary value of GenAI did not reside in providing correct answers. Rather, it functioned as an interacting agent that encouraged students to articulate and develop their own reasoning. This interpretation, however, must be qualified by the theoretical risk of algorithmic conformity. As reported in "Student Perspectives" section, a subset of students demonstrated revision patterns that progressively aligned with the AI platform's structural features and terminology. If students systematically revised their analyses to match AI feedback patterns rather than to express independently developed reasoning, the observable improvement in written work may partially reflect strategic platform alignment rather than genuine cognitive development. The current data cannot fully distinguish between these two possibilities at the individual level. This distinction has significant implications for how the scaffolding metaphor should be applied. Effective scaffolding is defined by its capacity to be progressively withdrawn as competence develops. If students learn to reason in ways that mirror AI logic rather than developing autonomous judgment, the scaffold may function as a cognitive prosthetic rather than a developmental support. Instructional designs should therefore explicitly train students to evaluate AI reasoning patterns, identify limitations in AI-generated analysis, and articulate the distinctions between their own reasoning and the AI's inferential framework.

The negative cases identified in the qualitative analysis introduce important qualifications. The surface-level use pattern suggests that active constructivist engagement is a condition requiring explicit instructional enforcement rather than a natural consequence of platform access. The resistance pattern among autonomy-oriented students suggests that the framework's applicability may be moderated by individual epistemological beliefs about legal reasoning development. The pre-consultation dependency pattern suggests that the staged scaffold-reduction logic may not have been sufficiently enforced at the individual level. Taken together, these findings indicate that the AI-Enhanced Legal Pedagogy Framework requires mechanisms for monitoring and differentiating individual GenAI engagement patterns.

The extended TAM demonstrated moderate-to-strong explanatory power ($R^2 = .59$ for behavioral intention, $R^2 = .52$ for attitude toward use). Core TAM relationships showed robust effects ($p < .001$), while contextual factors such as instructor guidance and prior experience showed significant but weaker influences ($p < .05$). This pattern indicates that student-centered perceptions outweigh external moderators in GenAI acceptance. The significant moderating role of instructor guidance ($\beta = .23$) highlights that technology acceptance in professional education is shaped by pedagogical relationships and implementation quality (Granić & Marangunić, 2019). As instructors shifted from information providers to learning facilitators, consistent with constructivism, GenAI enabled greater time for higher-order tasks including metacognitive skills training and complex reasoning modeling (Borge et al., 2024).

This study offers a preliminary conceptual integration of constructivism, TAM, and TPACK as applied to GenAI-enhanced legal education. It provides an organizing structure for future inquiry rather than a novel theoretical model. On the practical side, successful GenAI integration requires faculty development, authentic performance-based assessment strategies, and clear usage guidelines (Khlaif et al., 2025). A central finding is that

instructor facilitation significantly influences learning outcomes. Technology serves as a tool within the instructor-student relationship rather than as an independent pedagogical agent. Its effectiveness depends on how instructors structure and facilitate student engagement (Wang et al., 2024). The limitations of the present design, including the absence of a delayed post-test and the reliance on self-reported control group AI exposure measures, are discussed in "Study limitations" section.

Limitations and future research

Study limitations

Several limitations should be acknowledged. The single-case study design constrains generalizability. Findings are drawn from third-year students at one Chinese institution studying property law with one GenAI platform. The quasi-experimental design with non-random assignment introduces threats to internal validity. Although baseline equivalence was established on measured variables, unobserved confounds cannot be ruled out. This concern is particularly consequential given the large between-group effect size (Cohen's $d = 1.17$), which is atypically large for a non-randomized educational intervention and may reflect a greater degree of confounding than a study reporting more modest effects.

The study lacks a delayed post-test. Outcome assessment was conducted at the immediate conclusion of the intervention while students retained access to iFlytek Spark. The design therefore does not support conclusions about long-term retention of legal reasoning gains or their transferability to conditions without AI support. From the perspective of cognitive load theory (Sweller et al., 1998), it remains unclear whether observed improvements represent internalized reasoning schemas or performance benefits enabled by ongoing AI access. This limitation is especially salient given that legal practice routinely requires independent reasoning under examination and oral argument conditions. The present findings should be interpreted as evidence of immediate post-intervention improvement rather than lasting competence development.

Control group AI exposure was assessed primarily through self-reported surveys, which are subject to social desirability bias. Classroom observations and performance-based assessments provided supplementary evidence consistent with minimal AI use. However, the absence of objective behavioral monitoring means that undisclosed AI use outside class cannot be definitively ruled out. This represents a structural constraint of naturalistic quasi-experimental designs in real institutional settings.

The use of a commercial third-party platform introduces data governance limitations. Although informed consent procedures and anonymization protocols were implemented, student-submitted content was processed under iFlytek's commercial terms of service. Complete exclusion of submissions from platform-level data retention cannot be guaranteed. Future research should consider institutionally hosted AI infrastructure to ensure full researcher control over data processing.

The SEM analysis, conducted with 180 participants and a cases-to-indicators ratio of 4.5:1, approaches the lower boundary of recommended thresholds for stable latent variable estimation. Path coefficients for weaker pathways (H5, H6, H7) should be interpreted with caution. Future replications should target samples of 250 to 300 participants. Finally, the qualitative component included only three instructors. Instructor-level findings should be treated as hypothesis-generating rather than confirmatory.

Future research directions

Future studies should address the limitations identified above through several methodological improvements. First, self-reported AI use measures should be replaced with objective behavioral monitoring systems. Platform-level analytics or institutional network access logs would provide timestamped records of AI engagement and substantially strengthen internal validity. Where such monitoring is unavailable, randomized controlled designs assigning students at the individual level would offer a more rigorous alternative to naturalistic quasi-experimental approaches.

Second, delayed post-test assessments should be administered at least three to six months after intervention completion. These assessments should be conducted under conditions that preclude AI assistance, such as timed handwritten analyses or unassisted oral examinations. Researchers may also consider withdrawal designs in which a subset of participants transitions to AI-free instruction after the initial intervention. This would allow within-study assessment of retention and atrophy effects without multi-year follow-up. Establishing the durability of GenAI-supported learning gains is a prerequisite for credible claims about the long-term value of AI integration in professional education (Hyde-Vaamonde & Keller, 2025).

Third, future research should extend across multiple institutions, legal domains, and jurisdictions. Examining GenAI integration in criminal law, constitutional law, and contract law would identify domain-specific concerns. Cross-national comparisons between civil law and common law systems would illuminate cultural and institutional factors in GenAI adoption.

Finally, methodological tools are needed to distinguish genuine reasoning internalization from algorithmic conformity. Process-oriented approaches examining revision trajectories across successive drafts, rather than evaluating final written products alone, would help identify whether iterative improvement reflects authentic cognitive development or strategic alignment with AI feedback patterns. Development of validated metrics for measuring student-AI convergence across revision cycles would represent a meaningful contribution to the broader field of AI-assisted learning research.

Conclusion

The findings from the study suggest that integrating GenAI within a theoretically grounded pedagogical system was associated with meaningful improvements in legal reasoning outcomes, offering promising but preliminary evidence for the value of GenAI in legal education. The observed between-group differences (Cohen's $d = 1.17$, interpreted with caution given the non-randomized design), moderate to moderate-to-strong technology acceptance ($R^2 = .59$), and mixed qualitative findings collectively provide preliminary empirical support for the potential of GenAI integration in the legal education sector. The AI-Enhanced Legal Pedagogy Framework offers a preliminary conceptual integration of constructivist learning theory, TPACK, and the Technology Acceptance Model, providing an organizing structure to guide future research and practice in this domain rather than constituting a validated theoretical model. Crucially, the impact of GenAI depends upon the guidance of the teacher. It should be noted, however, that the absence of a delayed post-test represents a critical boundary on the conclusions that can be drawn from this study. Whether the



Fig. 8 Authentic iFlytek Spark interface showing a representative student-AI Socratic dialogue interaction during Phase 2 (Weeks 5–8). Screenshot captured from the iFlytek Spark web application using a hypothetical scenario constructed for pedagogical purposes. Interface in Mandarin Chinese. The exchange illustrates the guided questioning pattern in which the platform prompts the student to evaluate competing claims and construct counterarguments rather than providing a direct answer

observed reasoning gains reflect genuine internalization of legal reasoning competencies or AI-supported performance enhancement that may attenuate following scaffold removal remains an open empirical question that future research must address before stronger practical recommendations can be made.

It demonstrates the evidence-based ways of legal education: step-wise scaffolded learning with AI support, direct instruction for assessing AI outputs, and revised assessment with a focus on performative tasks rather than GenAI simulation. It suggests to the administrative and political audience a focus on professional development for faculty in tech, pedagogy, and content expertise; policies for ethics and integrity; and support for implementation to ensure consistency in quality. Looking ahead to a developing GenAI, there should be a thoughtful and evidence-based incorporation of AI in legal education, allowing for the use of human judgment and applying GenAI to provide scaffolded cognitive support in higher education to improve, not diminish, intellectual rigor, critical thinking, and professional skills in training competent, ethical, and analytical lawyers in a tech-infused environment.

Appendix A: Sample legal reasoning ability test (LRAT) ITEM

A.1 Item information

Item attribute	Details
Item number	LRAT-17 (Post-test form)
Legal domain	Adverse possession and bona fide acquisition
Reasoning category	Analytic reasoning
Difficulty index (pilot)	.58
Item-total correlation (pilot)	.61
Maximum score	20 points

A.2 Scenario prompt

Read the following scenario carefully and provide a written legal analysis responding to the questions below.

In 2008, Wang Fang purchased a residential property in a mid-sized Chinese city and held registered title. Due to employment relocation, Wang Fang left the property vacant and did not return. Beginning in 2010, Li Ming, a neighbor, began occupying the property, maintaining it, paying utility bills, and using it as his primary residence without Wang Fang's knowledge or consent.

In 2022, Wang Fang decided to sell the property and entered into a purchase agreement with Zhang Wei. Zhang Wei conducted a site inspection, observed Li Ming in occupation, but proceeded with the transaction after Wang Fang assured him that Li Ming was a temporary caretaker. Zhang Wei paid full market value and the transaction was registered in his name in March 2023.

Upon learning of the registration, Li Ming filed a claim asserting rights to the property. Wang Fang has left the country and cannot be located. Zhang Wei insists his registered title is valid and demands Li Ming vacate the property.

Question 1: Identify all legally relevant issues raised by this scenario under Chinese property law. (4 points).

Question 2: Apply the relevant legal rules and statutory provisions to each party's position, explaining how the law governs each claim. (8 points).

Question 3: Evaluate the competing claims and construct a reasoned legal argument identifying which party has the strongest legal position, acknowledging the weaknesses of that position. (8 points).

A.3 Scoring rubric (highest and lowest anchors)

Question	Score	Descriptor
Q1 (4 pts)	4	Identifies all four issues (adverse possession claim, bona fide purchaser defense, effect of registration on competing rights, relevance of prior knowledge of occupation) with precision and legal terminology
Q1	0	Fails to identify any legally relevant issue, or response is entirely non-responsive
Q2 (8 pts)	7–8	Accurately identifies and applies Articles 209, 311, and 312 of the Civil Code. Correctly explains bona fide acquisition conditions (reasonable price, good faith, completed registration) and applies each to the scenario facts. Addresses Li Ming's occupancy with doctrinal accuracy
Q2	0	No legally recognizable rule application, or reproduces statutory text without application to facts
Q3 (8 pts)	7–8	Constructs a clearly reasoned argument grounded in bona fide acquisition doctrine and registration primacy. Explicitly acknowledges the weakness arising from Zhang Wei's prior knowledge of occupation. Argument is logically structured and demonstrates awareness of legal ambiguity
Q3	0	No recognizable legal argument constructed

A.4 Annotated anchor responses (Question 3)

High-scoring response (8/8):

> "Under Chinese property law, Zhang Wei holds the strongest legal position by virtue of his completed registration under Article 209 of the Civil Code, which establishes that property rights are created, modified, or extinguished upon registration. The bona fide acquisition doctrine under Article 311 further protects Zhang Wei's title provided that: the transaction was at a reasonable price, Zhang Wei acted in good faith, and registration was completed. The first and third conditions are clearly satisfied. The critical issue is whether Zhang Wei's prior knowledge of Li Ming's occupation defeats the good faith requirement. A court may find that awareness of a visible occupant should have prompted further inquiry, potentially negating good faith. Nevertheless, absent a judicial determination, the registered title presumption favors Zhang Wei. Li Ming's occupancy does not confer registered title under Chinese civil law, and adverse possession does not operate in the same form as in common law jurisdictions. Zhang Wei's position is strongest, though not without meaningful legal risk on the good faith question."

This response demonstrates statutory application with article citation, multi-perspective evaluation, acknowledgment of weakness, and awareness of civil law distinctions.

Low-scoring response (2/8):

"Zhang Wei should win because he paid for the property and registered it. Registration is very important in Chinese property law. Li Ming did not have a contract and did not register the property so he cannot claim ownership. Wang Fang sold the property legally so Zhang Wei is the rightful owner."

This response reproduces a general principle without statutory grounding, fails to address bona fide acquisition conditions or the good faith issue, and contains no acknowledgment of competing claims.

Note The parallel pre-test item addresses a co-ownership dispute of equivalent cognitive demand and difficulty index (difference = .03). Scoring was conducted blind to group assignment.

Appendix B: Representative GenAI prompt examples and instructor scaffolding protocols

Figure 8 presents an authentic iFlytek Spark interface screenshot capturing a representative student-AI interaction during Phase 2 (Weeks 5–8), illustrating the Socratic dialogue pattern in which the platform guided students to analyze competing claims and develop counterarguments.

B.1 Overview

The following examples illustrate representative prompt structures used by students and instructors during the 16-week intervention. Prompts are organized by intervention phase (Table 1) and GenAI role. All case scenarios are hypothetical and were constructed specifically for pedagogical purposes. Prompts were originally administered in Mandarin Chinese; English translations are provided here for reporting purposes.

B.2 Phase 1 (weeks 1–4): information provider and reasoning demonstrator

Learning objective: Identify property law issues and apply basic ownership principles.

Student prompt example 1 (issue identification):

The following scenario involves a dispute between two neighbors over a boundary fence. Please identify the key legal issues under Chinese property law and explain which statutory provisions are most relevant: [scenario text].

Student prompt example 2 (concept clarification):

Please explain the concept of bona fide acquisition under Chinese property law, including the conditions that must be satisfied and a brief example of how it applies in practice.

Instructor scaffolding protocol (Phase 1):

Before AI interaction: instructor introduces the case scenario and asks students to attempt independent issue identification for 10 minutes

During AI interaction: students submit their own analysis first, then query iFlytek Spark for comparison

After AI interaction: instructor facilitates whole-class discussion on discrepancies between student analyses and AI-generated responses, emphasizing that AI outputs require critical evaluation against course materials

Fidelity check: instructors recorded whether the pre-AI independent attempt step was completed for each session

B.3 Phase 2 (weeks 5–8): socratic dialogue partner

Learning objective: Analyze complex property rights conflicts and evaluate competing claims.

Student prompt example 1 (multi-party dispute):

> In the following adverse possession scenario, Party A has occupied Party B's registered land for 18 years without interruption. Party C subsequently purchases the land from Party B in good faith. Please identify the competing legal claims of each party and explain the reasoning behind each position, including any weaknesses.

Student prompt example 2 (Socratic follow-up):

> You argued that Party A's claim would likely succeed. What is the strongest counterargument that Party B's attorney could make? What evidence would be most decisive in resolving this dispute?

Instructor scaffolding protocol (Phase 2):

Before AI interaction: instructor models a Socratic dialogue sequence with iFlytek Spark using a different scenario, demonstrating how to probe AI responses with follow-up questions

Usage guideline provided to students: *"Do not accept the first response. Follow up with at least two probing questions that challenge or extend the AI's initial analysis."*

After AI interaction: students submit a written reflection identifying one point where the AI's reasoning was incomplete or required correction

Fidelity check: interaction logs reviewed to confirm minimum two-exchange dialogue sequences per session

B.4 Phase 3 (weeks 9–12): writing assistant and feedback provider

Learning objective: Construct legal arguments and develop persuasive legal writing.

Student prompt example 1 (draft feedback):

> The following is my draft legal opinion arguing that the plaintiff should succeed in recovering the disputed property. Please evaluate the logical structure of my argument, identify any gaps in my application of legal principles to the facts, and suggest how the persuasiveness of the conclusion could be strengthened: [draft text].

Student prompt example 2 (counterargument development):

> Based on the legal opinion I, submitted, please construct the strongest possible counterargument that the defendant's counsel could raise, citing relevant statutory provisions and precedents under Chinese property law.

Instructor scaffolding protocol (Phase 3):

Before AI interaction: instructor provides a marking rubric specifying the criteria against which AI feedback should be evaluated

Usge guideline provided to students: "Use AI feedback as one input among several. You must also apply the course rubric independently and explain in your revision notes which AI suggestions you accepted, which you rejected, and why."

After AI interaction: peer review session in which students compare AI-generated feedback with feedback from a classmate

Fidelity check: revision notes, submitted alongside final drafts to document student reasoning about AI feedback integration

B.5 Phase 4 (weeks 13–16): strategic planning assistant and opposing counsel simulator

Learning objective: Synthesize property law knowledge and prepare oral advocacy.

Student prompt example 1 (moot court preparation):

> I am preparing to argue the appellant's position in the following property dispute. Please identify the three strongest arguments available to the appellant, the most likely objections the respondent will raise, and the key statutory provisions and precedents I should be prepared to address in oral argument: [case summary].

Student prompt example 2 (opposing counsel simulation):

> Please take the role of opposing counsel in this moot court case and cross-examine my opening argument as rigorously as possible, identifying any logical inconsistencies, unsupported factual assumptions, or misapplications of legal principle: [argument text].

Instructor scaffolding protocol (Phase 4):

Before AI interaction: instructor specifies that AI preparation is supplementary and that oral performance will be assessed on real-time reasoning independent of prepared materials

Usage guideline provided to students: *"AI can help you anticipate questions but cannot answer them for you in the oral argument. Your goal is to internalize the reasoning, not to memorize AI-generated responses."*

After AI interaction: mock oral argument conducted under timed conditions without AI access, assessed by instructor using the standard moot court rubric

Fidelity check: comparison of AI-prepared arguments with actual oral performance to assess degree of independent reasoning demonstrated

B.6 Instructor preparation protocol summary

Prior to the intervention, all three experimental group instructors completed a structured preparation protocol addressing each TPACK intersection:

Training component	Focus	Duration
Technological knowledge (TK)	iFlytek Spark interface, capability boundaries, hallucination patterns in legal contexts	3 h
Technological pedagogical knowledge (TPK)	Designing AI-mediated Socratic sequences; structuring pre/during/post AI interaction phases	4 h
Technological content knowledge (TCK)	Identifying and correcting AI errors in property law doctrine; jurisdictional limitation awareness	3 h
Integrated TPACK	Designing activities requiring reasoning that exceeds AI capability; fidelity monitoring procedures	2 h

Instructors met weekly during the intervention to review usage logs, discuss implementation challenges, and calibrate scaffolding decisions across sections.

Note All prompts were administered via the iFlytek Spark platform interface in Mandarin Chinese. Scenarios were hypothetical and constructed by the research team. Prompt structures were provided to students as guidelines rather than fixed scripts, allowing adaptation to individual case analyses.

Funding This research received grants from Provincial Curriculum Ideological and Political Model Curriculum Property Law. (11230—03342022005).

Data Availability The data supporting the findings of this study are largely included within the article and its appendices. Additional raw data are available from the corresponding author upon reasonable request, subject to approval from the relevant institutional authorities.

Declarations

Conflict of interest The authors declare no conflicts of interest.

Ethical approval Research ethics approval was obtained by the Institutional Review Board of Hefei University of Technology.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Abrami, P. C., Cohen, P. A., & d'Apollonia, S. (1988). Implementation problems in meta-analysis. *Review of Educational Research, 58*(2), 151–179.

Al-Adwan, A. S., Li, N., Al-Adwan, A., Abbasi, G. A., Albelbisi, N. A., & Habibi, A. (2023). Extending the technology acceptance model (TAM) to predict university students' intentions to use metaverse-based learning platforms. *Education and Information Technologies, 28*(11), 15381–15413.

Allsop, D. B., Chelladurai, J. M., Kimball, E. R., Marks, L. D., & Hendricks, J. J. (2022). Qualitative methods with NVivo software: A practical guide for analyzing qualitative data. *Psych, 4*(2), 142–159.

Almulla, M. A. (2023). Constructivism learning theory: A paradigm for students' critical thinking, creativity, and problem solving to affect academic performance in higher education. *Cogent Education, 10*(1), Article 2172929.

Alshammari, S. H., & Rosli, M. S. (2020). A review of technology acceptance models and theories. *Innovative Teaching and Learning Journal, 4*(2), 12–22.

Andreeva, D., & Savova, G. (2024). Artificial intelligence in the legal field: Law students' perspective. arXiv:2410.09937.

Bobula, M. (2024). Generative artificial intelligence (AI) in higher education: A comprehensive review of challenges, opportunities, and implications. *Journal of Learning Development in Higher Education*. https://doi.org/10.47408/jldhe.vi30.1137

Borge, M., Smith, B., & Aldemir, T. (2024). Using generative AI as a simulation to support higher-order thinking. *International Journal of Computer-Supported Collaborative Learning, 19*(4), 479–532.

Chan, C. K. Y., & Colloton, T. (2024). *Generative AI in higher education: The ChatGPT effect*. Taylor & Francis.

Chee, H., Ahn, S., & Lee, J. (2025). A competency framework for AI literacy: Variations by different learner groups and an implied learning pathway. *British Journal of Educational Technology, 56*(5), 2146–2182.

Cheung, R., & Vogel, D. (2013). Predicting user acceptance of collaborative technologies: An extension of the technology acceptance model for e-learning. *Computers & Education, 63*, 160–175.

Chimbwanda, V. (2022). *Embedding skills in African customary law and culture in the LLB curriculum: An empirical study of pedagogical approaches in selected African university law schools* [Doctoral thesis, School of Advanced Study, University of London].

Clark, R. E. (1983). Reconsidering research on learning from media. *Review of Educational Research, 53*(4), 445–459.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319–340.

Dawson, P. (2020). *Defending assessment security in a digital world: Preventing e-cheating and supporting academic integrity in higher education*. Routledge.

Demchenko, M. V., Gulieva, M. E., Larina, T. V., & Simaeva, E. P. (2021). Digital transformation of legal education: Problems, risks and prospects. *European Journal of Contemporary Education, 10*(2), 297–307.

Do, H.-N., Do, B. N., & Nguyen, M. H. (2023). How do constructivism learning environments generate better motivation and learning strategies? The design science approach. *Heliyon*. https://doi.org/10.1016/j.heliyon.2023.e22862

Draucker, C. B., Rawl, S. M., Vode, E., & Carter-Harris, L. (2020). Integration through connecting in explanatory sequential mixed method studies. *Western Journal of Nursing Research, 42*(12), 1137–1147.

Garingan, D., & Pickard, A. J. (2021). Artificial intelligence in legal practice: Exploring theoretical frameworks for algorithmic literacy in the legal information profession. *Legal Information Management, 21*(2), 97–117.

Granić, A. (2022). Educational technology adoption: A systematic review. *Education and Information Technologies, 27*(7), 9725–9744.

Granić, A., & Marangunić, N. (2019). Technology acceptance model in educational context: A systematic literature review. *British Journal of Educational Technology, 50*(5), 2572–2593.

Grassini, S. (2023). Shaping the future of education: Exploring the potential and consequences of AI and ChatGPT in educational settings. *Education Sciences, 13*(7), 692.

Guo, M. (2024). The influence of educational and emotional support on e-learning continuance intention: A two-stage approach PLS-SEM and ANN analysis. *SAGE Open, 14*(4), Article 21582440241280793.

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.

Han, J.-H., & Sa, H. J. (2022). Acceptance of and satisfaction with online educational classes through the technology acceptance model (TAM): The COVID-19 situation in Korea. *Asia Pacific Education Review, 23*(3), 403–415.

Hyde-Vaamonde, C., & Keller, A. (2025). Testing the frontier: Generative AI in legal education and beyond. *Law, Technology and Humans, 7*(3), 52–61.

Jiménez Sierra, Á. A., Ortega Iglesias, J. M., Cabero-Almenara, J., & Palacios-Rodríguez, A. (2023). Development of the teacher's technological pedagogical content knowledge (TPACK) from the lesson study: A systematic review. *Frontiers in Education, 8*, Article 1078913.

Khlaif, Z. N., Alkouk, W. A., Salama, N., & Abu Eideh, B. (2025). Redesigning assessments for AI-enhanced learning: A framework for educators in the generative AI era. *Education Sciences, 15*(2), Article 174.

Kholid, M. N., Hendriyanto, A., Sahara, S., Muhaimin, L. H., Juandi, D., Sujadi, I., & Adnan, M. (2023). A systematic literature review of technological, pedagogical and content knowledge (TPACK) in mathematics education: Future challenges for educational practice and research. *Cogent Education, 10*(2), Article 2269047.

Kim, M., & Adlof, L. (2024). Adapting to the future: ChatGPT as a means for supporting constructivist learning environments. *TechTrends, 68*(1), 37–46. https://doi.org/10.1007/s11528-023-00899-x

Lee, S. S., & Moore, R. L. (2024). Harnessing generative AI (GenAI) for automated feedback in higher education: A systematic review. *Online Learning, 28*(3), 82–106.

Leijen, Ä., Malva, L., Pedaste, M., & Mikser, R. (2022). What constitutes teachers' general pedagogical knowledge and how it can be assessed: A literature review. *Teachers and Teaching, 28*(2), 206–225.

Li, W. (2022). Envisioning the role of educators' technological pedagogical and content knowledge and self-regulated learning in an English as a foreign language context. *Frontiers in Psychology, 13*, Article 943072.

Lin, Y., & Yu, Z. (2023). Extending technology acceptance model to higher-education students' use of digital academic reading tools on computers. *International Journal of Educational Technology in Higher Education, 20*(1), 34.

Love, H. R., Cook, B. G., & Cook, L. (2022). Mixed-methods approaches in special education research. *Learning Disabilities Research & Practice, 37*(4), 314–323.

Mertz, E. (2007). *The language of law school: Learning to "think like a lawyer."* Oxford University Press.

Mishra, N. R. (2023). Constructivist approach to learning: An analysis of pedagogical models of social constructivist learning theory. *Journal of Research and Development, 6*(1), 22–29.

Mulyani, H., Istiaq, M. A., Shauki, E. R., Kurniati, F., & Arlinda, H. (2025). Transforming education: Exploring the influence of generative AI on teaching performance. *Cogent Education, 12*(1), Article 2448066.

Neumann, R. K., Margolis, E., & Stanchi, K. M. (2021). *Legal reasoning and legal writing*. Wolters Kluwer.

Paskins, A. (2023). *Classroom teachers' perceptions of their knowledge of content and pedagogy in terms of a technology-driven lesson* [Doctoral dissertation, Walden University].

Rahimi, M. (2025). Synergising dialogic teaching with competencies-trained GenAI dialoguing for critical thinking and communication competencies. *Journal of University Teaching & Learning Practice, 22*(2), 1–22.

Ringle, C. M., Wende, S., & Becker, J.-M. (2015). *SmartPLS 3* [Computer software]. SmartPLS. https://www.smartpls.com

Shaw, G. (2012). A heretical view of teaching: A contrarian looks at teaching, the Carnegie report, and best practices. *Touro Law Review, 28*, 1239.

Stuckey, R. (2009). Best practices or not, it is time to re-think legal education. *Clinical Law Review, 16*, 307.

Sullivan, W. M., Colby, A., Wegner, J. W., Bond, L., & Shulman, L. S. (2007). *Educating lawyers: Preparation for the profession of law*. John Wiley & Sons.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Sweller, J., Van Merriënboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), 251–296.

Teo, T. (2009). The impact of subjective norm and facilitating conditions on pre-service teachers' attitude toward computer use: A structural equation modeling of an extended technology acceptance model. *Journal of Educational Computing Research, 40*(1), 89–109.

Wang, H., Dang, A., Wu, Z., & Mac, S. (2024). Generative AI in higher education: Seeing ChatGPT through universities' policies, resources, and guidelines. *Computers & Education: Artificial Intelligence, 7*, Article 100326.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Haifang Liu is a faculty member at the School of Humanities and Law, Hefei University of Technology, Hefei, Anhui, China. Her research interests include legal education, curriculum development, and the integration of technology in law teaching. She is the principal investigator of the Provincial Curriculum Ideological and Political Model Curriculum Project on Property Law (Grant No. 11230—03342022005). Her current work focuses on faculty role transformation and professional identity in AI-enhanced legal education contexts.