Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Generative AI Use and Doctoral Research Satisfaction: The Nonlinear Roles of Supportive and Substitutive Use

Yaozheng Zhi

20230116@sxnu.edu.cn

Shanxi Normal University

Wei Yang

Taiyuan University of Science and Technology

Kaili Huang

Huazhong University of Science and Technology

Yexin Li

Huazhong University of Science and Technology

Article

Keywords: supportive and substitutive use, doctoral students, research satisfaction

Posted Date: November 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11328922/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Generative AI Use and Doctoral Research Satisfaction: The Nonlinear Roles of Supportive and Substitutive Use

Abstract: Generative AI is profoundly reshaping the research ecosystem for doctoral students, yet how use is associated with their research satisfaction remains unclear. Drawing on cognitive load theory and self-determination theory, this study used data from a questionnaire survey of 786 Chinese doctoral students to distinguish between supportive and substitutive use and to examine their associations with doctoral students' research satisfaction. The results showed that both supportive and substitutive use had significant inverted U-shaped associations with doctoral research satisfaction. Moderate frequencies of use were associated with higher satisfaction, whereas excessively high frequencies were associated with lower satisfaction. Heterogeneity analysis indicated that the inverted U-shaped association held across STEM, medical and health, and humanities and social sciences disciplines, with the estimated peak being highest in STEM disciplines. AI efficiency perception showed a positive conditional instantaneous indirect effect in the association between AI use and research satisfaction, whereas AI risk perception showed a negative one. Overtime work pressure strengthened the quadratic associations between AI use and both AI efficiency perception and AI risk perception.

Keywords: supportive and substitutive use; doctoral students; research satisfaction

Introduction

The groundbreaking advancement of generative AI is catalyzing a profound intelligent transformation within higher education, with particularly significant repercussions for doctoral education. As a pivotal phase in the cultivation of high-caliber innovative talent, the quality of doctoral training directly affects national scientific and technological competitiveness and the efficiency of knowledge production. Currently, generative AI has been extensively adopted across key facets of doctoral research practice, encompassing literature retrieval, data processing, code development, and academic writing, thereby offering novel technological avenues for enhancing research productivity (Gao et al., 2024). However, while generative AI enhances doctoral research efficiency, it may also introduce concerns such as cognitive dependence and skill erosion. Greater technical efficiency does not necessarily translate into a better research experience. According to Nature's 2025 Global PhD Survey, more than half of doctoral students use generative AI tools weekly in their research, a figure that highlights the growing pervasiveness of AI in doctoral research practice. Yet how generative AI use relates to doctoral research satisfaction remains unsettled in the literature.

Doctoral research satisfaction reflects doctoral students' sense of achievement, fulfillment, and professional identity throughout the research process, and serves as one of the key dimensions for evaluating the quality of doctoral training (Khan & Nasim, 2025). Conceptually, this construct is multidimensional: it encompasses not only doctoral students' perceptions of the research content itself, but also their evaluations of the external conditions and support systems that enable research

activities to proceed (Sverdlik et al., 2018). Regarding the relationship between generative AI and doctoral research satisfaction, two competing perspectives have emerged in the literature.

The supportive view argues that generative AI can significantly enhance research satisfaction, operating through three distinct mechanisms. At the level of efficiency gains, generative AI can rapidly complete routine tasks such as data cleaning and formatting, substantially shortening the preparatory phase of research (Kasneci et al., 2023). The subjective well-being of knowledge workers is closely tied to their perception of work efficiency. By compressing time spent on low-value tasks, generative AI indirectly strengthens doctoral students' perceived control over their research progress and their sense of accomplishment (Alireza et al., 2026; Khalifa & Albadawy, 2024). At the level of cognitive optimization, generative AI functions as a "cognitive collaborator" in the research process, assisting doctoral students with literature classification, idea synthesis, and structural organization (Rahman & Watanobe, 2023). It alleviates the cognitive burden associated with information filtering and preliminary integration, thereby freeing up limited attentional resources. This may allow doctoral students to redirect their cognitive efforts toward higher-order academic activities, such as theoretical construction and logical reasoning (Li et al., 2024; Andersen et al., 2025). At the level of capability expansion, generative AI helps doctoral students transcend disciplinary boundaries and technical barriers, facilitating rapid acquisition of foundational knowledge and analytical methods across fields. It broadens their research horizons and diversifies their channels for knowledge acquisition (Zou & Huang, 2024; Dai et al., 2023). This allows doctoral students to grasp the intellectual contours of unfamiliar domains in a shorter time frame, thereby bolstering their research confidence and initiative in scholarly exploration.

The critical view argues that the use of generative AI may negatively affect doctoral research satisfaction. This concern manifests in three primary aspects. Generative AI may lead to cognitive displacement. The ease of use associated with generative AI may tempt doctoral students to delegate critical cognitive tasks to AI systems, thereby bypassing essential steps that should constitute their own intellectual training (Uludağ, 2025). When these core activities are consistently outsourced, doctoral students are likely to experience heightened anxiety about the erosion of their own cognitive capabilities, which in turn may diminish their research satisfaction (He et al., 2026). Furthermore, generative AI may also result in skill degradation. Over-reliance on generative AI for core research activities such as data analysis, code writing, and text generation may lead to a gradual decline in doctoral students' scholarly judgment and foundational research competencies. Consequently, their ability to conduct independent research may not receive adequate cultivation through hands-on practice (Zhu & Yang, 2026; Ringo, 2025). In addition, generative AI raises concerns about academic ethics. The inherent uncertainty of AI-generated outputs introduces novel challenges to doctoral students regarding the integrity of the research process, the authenticity of data, and the assurance of academic honesty (Zhou et al., 2024). If AI-generated content is used inappropriately or incorporated into the research without proper verification, it may constitute a form of implicit academic misconduct (Rice et al., 2024). When doctoral students begin to question whether their research can withstand the scrutiny of academic standards, their research satisfaction is likely to decline accordingly.

The literature reviewed above reveals multiple pathways through which generative AI may relate to doctoral research satisfaction, offering valuable reference points for understanding how

technological intervention shapes doctoral research experiences. Nevertheless, several gaps remain to be addressed. First, existing discussions have largely remained at the level of theoretical reasoning, with limited empirical testing based on large-scale data. Although both the positive and critical perspectives are grounded in reasonable rationales, their conclusions are mostly derived from localized observations, which limits their generalizability and calls for further verification. Second, most empirical studies treat generative AI use as a single overarching construct, without distinguishing between different modes of interaction. Whether AI serves as an assistive collaborator or replaces independent intellectual effort may produce fundamentally different effects, yet merging heterogeneous AI use behaviors into a single indicator can obscure these differential effects and lead to inconsistent conclusions. Third, existing studies largely assume a linear relationship between AI use and research outcomes, leaving nonlinear associations insufficiently examined and potentially overlooking the distinct effects of moderate versus excessive use. Finally, doctoral students may experience both efficiency and risk perceptions as they use generative AI in their research. These two psychological responses may shape research satisfaction, yet their roles in the relationship between AI use and research satisfaction remain systematically unexamined. Overtime work pressure is a prevalent phenomenon in doctoral research environments. This contextual factor may also shape the strength of these associations, yet its specific role awaits further investigation.

This study draws on cognitive load theory and self-determination theory as its analytical framework and uses data from a questionnaire on Chinese doctoral students. It distinguishes between two dimensions of generative AI use: supportive use and substitutive use. Three questions are examined: whether these two types of use show nonlinear associations with doctoral research satisfaction; whether AI efficiency perception and AI risk perception mediate the relationship between AI use and doctoral research satisfaction; and whether overtime work pressure moderates the associations between AI use and both AI efficiency perception and AI risk perception.

This study makes two theoretical contributions. First, it integrates cognitive load theory and self-determination theory into a unified analytical framework that links individual cognitive resource reallocation to psychological need experiences. Cognitive load theory is well suited to explaining how AI involvement reshapes the balance among intrinsic, extraneous, and germane loads, but it does not adequately address how these cognitive changes translate into doctoral students' subjective experiences. Self-determination theory fills this gap by explaining how autonomy and competence are enhanced or weakened during AI use. By incorporating overtime work pressure as an institutional condition, the framework connects individual-level mechanisms with the social and organizational context in which AI use occurs. Second, it distinguishes supportive use from substitutive use, and examines their nonlinear associations with research satisfaction. Existing research has largely treated AI use as a single overarching construct. Through theoretical analysis and empirical validation, the distinction between the two types of use is clarified in terms of their cognitive load reconfiguration paths and psychological need mechanisms, providing an operational and testable classification framework for future social research on AI use in academic work.

Theoretical Foundations and Research Hypotheses

Direct Associations Between Generative AI and Doctoral Research Satisfaction

Cognitive load theory provides an important theoretical framework for understanding the constraints of information processing in complex cognitive tasks. This theory posits that individuals have limited working memory capacity, and when the information processing demands of a task exceed this capacity, both learning outcomes and problem-solving abilities are impaired (Chen et al., 2023). Within this framework, cognitive load comprises three distinct types. Intrinsic load is determined by the complexity of the task itself, which depends on the number of information elements that must be processed simultaneously and the complexity of their interactions, as exemplified by tasks such as variable deduction and multifactor reasoning. Extraneous load is influenced by the manner in which information is presented and environmental distractions, including the additional cognitive burden imposed by literature retrieval and formatting adjustments. Germane load reflects the cognitive effort that individuals actively invest in schema construction and the automation of cognitive processing (Kirschner et al., 2018). Doctoral research activities, including literature review, theoretical construction, and data analysis, are characterized by high cognitive complexity, rendering them particularly susceptible to cognitive overload (Sverdlik et al., 2018). The introduction of generative AI may not simply increase or decrease the total amount of cognitive load. Rather, it may reshape the composition and balance among the three types of loads. Whether this restructuring is empowering or burdening depends on the specific mode of AI use adopted by doctoral students.

Cognitive load theory primarily explains the relationship of generative AI on doctoral students' research experiences from the perspective of cognitive resource allocation. Yet it does not adequately address how psychological need experiences, such as autonomy and competence, shape research satisfaction during AI use. Self-determination theory provides a complementary framework (Deci & Ryan, 2000). The theory posits that individuals sustained positive development and subjective well-being depend on the satisfaction of three basic psychological needs: autonomy, competence, and relatedness. By focusing on how individuals experience psychological need fulfillment in their interactions with the environment, it offers a complementary perspective for understanding the psychological pathways through which AI use relates to doctoral research satisfaction.

Existing research has confirmed that the effects of generative AI on doctoral students' research work vary by mode of use. Drawing on the classification proposed by Guo and Zhu (2025), this study distinguishes the two dimensions of AI use: supportive use and substitutive use. Supportive use refers to doctoral students' employment of generative AI as a cognitive collaborator in higher-order academic tasks, such as research design optimization, experimental planning, and literature tracking. Students retain overall control and final decision-making authority over research questions, while AI assists with information retrieval, preliminary organization, and language polishing. Substitutive use refers to doctoral students' reliance on generative AI as a task executor to directly produce substantive academic outputs, including data analysis, manuscript drafting, and code debugging. Although these tasks differ in cognitive complexity and specific content, they share a common feature: AI directly intervenes in core research production, replacing processes that would otherwise require doctoral students' independent intellectual effort.

Supportive use primarily involves the reconfiguration of extraneous and germane cognitive load. At moderate levels, AI engagement in rule-based tasks such as literature retrieval and language polishing effectively reduces extraneous cognitive load and frees limited working memory resources (Kim et al., 2025). This enables doctoral students to redirect their cognitive efforts toward higher-order academic activities, such as theoretical construction and logical reasoning. Students gain a stronger sense of autonomy and confidence in their academic work, both of which are associated with higher research satisfaction (Baber et al., 2024). However, this positive association does not hold indefinitely. When supportive use exceeds a certain threshold, the role of AI shifts from "cognitive assistance" to "substitute for thinking". Doctoral students no longer need to perform critical reasoning steps such as literature critique and theoretical integration themselves, but instead directly accept AI-generated outcomes (Pereira et al., 2024). This transition turns students from active cognitive processors into passive recipients. As cognitive engagement decreases, confidence in academic judgment and sense of competence may weaken, which may in turn be associated with lower research satisfaction (Lim et al., 2023).

Substitutive use similarly reshapes how cognitive load is experienced. When AI takes over certain task components that carry intrinsic cognitive load, the number of information elements students need to process themselves may decrease, which may free cognitive resources for higher-level research engagement. With these resources released, students may feel more in control of their research progress and more competent in completing tasks, experiences that correspond to higher research satisfaction (Ray, 2023). However, this positive association does not hold indefinitely. When AI involvement in core research production exceeds a moderate level, its role shifts from task executor to cognitive substitute. Students no longer engage deeply in complex cognitive processes but instead directly adopt AI-generated outputs (Zhou & Zhang, 2024). The shift may move them from active executors of research tasks to passive reviewers of AI products, reducing opportunities for cognitive engagement and limiting the accumulation of research skills (Meng et al., 2025). Students may experience anxiety and self-doubt when they perceive their independent research capabilities declining due to AI reliance, which may correspond to lower research satisfaction (Guleria et al., 2023).

The foregoing analysis suggests that both supportive and substitutive use may exhibit nonlinear associations with doctoral research satisfaction. Accordingly, the following hypotheses are proposed:

H1: Supportive use shows an inverted U-shaped association with doctoral research satisfaction. Moderate levels of supportive use are associated with higher research satisfaction, whereas excessive levels correspond to lower satisfaction.

H2: Substitutive use shows an inverted U-shaped association with doctoral research satisfaction. Moderate levels of substitutive use are associated with higher research satisfaction, whereas excessive levels correspond to lower satisfaction.

Mechanisms of the Associations Between Generative AI and Doctoral Research Satisfaction

Both supportive and substitutive use may exhibit nonlinear associations with doctoral research satisfaction. However, the psychological pathways through which this relationship operates remain

unclear. To address this gap, the present study further examines the mediating and moderating mechanisms.

Mediating Mechanisms: AI Efficiency Perception and AI Risk Perception

AI efficiency perception captures doctoral students' subjective experience of cognitive load reconfiguration (Khan & Ullah, 2024). When students perceive that generative AI helps reduce cognitive investment in literature retrieval, data organization, and language polishing, they are able to shift their cognitive focus toward core research activities such as theoretical construction and logical reasoning. As cognitive resources are released, students may gain a stronger sense of control over the research process and greater confidence in their task competence (Lo et al., 2024). These experiences correspond to more positive research satisfaction. However, when AI use exceeds a moderate range, the additional efficiency gains from AI use begin to diminish. Students find it harder to sustain the gains in autonomy and competence that come from task progression, and the room for further improvement in efficiency perception becomes saturated. Under such conditions, further increases in AI use may yield smaller gains in perceived efficiency, which may weaken the positive indirect association between AI use and research satisfaction through efficiency perception.

AI risk perception stems from doctoral students' concerns about generative AI replacing their own cognitive operations (Mohammad et al., 2026). When AI intervenes in core cognitive processes, students are prone to bypass necessary intellectual training and lose opportunities to deepen their understanding through active cognitive engagement, thereby developing doubts about their own abilities (Zou & Zhao, 2025). Such doubts consume cognitive resources, weaken deep processing capacity, and undermine their sense of autonomy and competence, which aligns with lower research satisfaction. As AI use grows, cognitive substitution expands in both scope and depth, giving rise to accumulating concerns about reduced cognitive engagement, uncertain AI outputs, and declining personal competence. These concerns may heighten AI risk perception, consume cognitive resources, weaken deep processing capacity, and undermine students' sense of autonomy and competence. Students' distrust of AI outputs further occupies cognitive space, allowing non-productive load to gradually take over and compress the resources originally available for schema construction and deep thinking. This process ultimately reinforces the negative association with research satisfaction (Safdar et al., 2024).

It is evident that AI efficiency perception and AI risk perception represent two important psychological pathways through which generative AI use relates to doctoral research satisfaction. Accordingly, the following hypotheses are proposed:

H3: AI efficiency perception positively mediates the relationship between generative AI use and doctoral research satisfaction, with this mediating effect varying across different frequencies of use.

H4: AI risk perception negatively mediates the relationship between generative AI use and doctoral research satisfaction, with this mediating effect varying across different frequencies of use.

Moderating Mechanism: Overtime Work Pressure

Overtime work pressure reflects the prevailing norms of extended working hours in doctoral students' institutions, including organizational expectations for nighttime and weekend work (Hill & Thabet, 2021). Under high time pressure, doctoral students' motivation for using generative AI shifts from pursuing deep understanding to seeking rapid output, with more attention directed

toward task completion rather than knowledge internalization (Van et al., 2024). This shift may not only be associated with a stronger positive linear association between AI use and efficiency perception, but may also change the rate at which efficiency perception varies with AI use. When students are oriented toward producing outputs quickly, the initial efficiency gains from AI may become more pronounced. At the same time, the saturation effect may arrive sooner, making the quadratic association between AI use and efficiency perception more pronounced under high time pressure.

In addition, high time pressure leads doctoral students to devote less effort to verifying and reflecting on AI-generated content, as they are driven by institutional expectations to produce more output (Stefano et al., 2025). Such reduced verification may not only strengthen the linear relationship between AI use and risk perception, but also accelerate the rate at which risk perception accumulates as AI use increases. Under high time pressure, concerns about skill degradation may compound more rapidly, making the quadratic association between AI use and risk perception more pronounced.

Based on this, the following hypothesis is proposed:

H5: Overtime work pressure positively moderates the quadratic associations between AI use and both AI efficiency perception and AI risk perception.

Methodology

Variables

Dependent Variable

Herzberg's (1959) Two-Factor Theory was adopted as the theoretical framework for measuring doctoral research satisfaction. The theory posited that an individual's job satisfaction was jointly determined by two categories of factors. Motivators related to the content of the work itself: a sense of achievement, recognition, challenging work, and personal growth. Their presence generated satisfaction. Hygiene factors related to the work environment, such as compensation, interpersonal relationships, organizational policies, and quality of supervision. Their absence led to dissatisfaction.

Research satisfaction was measured using twelve items developed on the basis of the Nature (2025) Global PhD Survey indicators (Q14a-1 to Q14a-12) and adapted to the research context of Chinese doctoral students. The wording of the items was adjusted to fit this context, without altering the core meaning of the original indicators. Drawing on this framework, the twelve items were classified into two categories. Motivators directly reflect doctoral students' intrinsic experiences in research activities, including academic autonomy, research guidance, academic participation, competence development, intellectual challenge, and academic community. Hygiene Factors represent the institutional and relational conditions that sustain research, including social networks, supervisory relationship, compensation, career guidance, work-life balance, and international opportunities. Although the hygiene factors do not directly constitute research content, they represent necessary conditions for a sustainable and satisfying research experience. The twelve items were theoretically classified into the two Herzberg categories, yet they jointly reflect an overall evaluative judgment of the doctoral research experience. Accordingly, research satisfaction was modeled as a single latent dimension. Latent scores for doctoral research satisfaction were then derived via an ordinal latent-variable model.

Core Independent Variables

The core independent variables were doctoral students' generative AI use, comprising two distinct latent constructs: supportive use and substitutive use. Supportive use items were drawn from the Nature (2025) Global PhD Survey Questionnaire (Q33-4 to Q33-6), covering experimental design, literature tracking, and literature synthesis. Given that language polishing is one of the most frequent scenarios in which doctoral students employ generative AI (Kim et al., 2025), one self-developed item on language polishing and refinement was added. Four items were ultimately retained to measure supportive use. Substitutive use items were adopted directly from the Nature (2025) Global PhD Survey Questionnaire (Q33-1, Q33-2, Q33-3, and Q33-7), covering the use of generative AI in peer review, article writing, data collection and analysis, and job application preparation.

Mediating and Moderating Variables

The two mediating variables were AI efficiency perception and AI risk perception. AI efficiency perception was measured by three items, covering time cost savings, practical research benefits, and subjective sense of research gain. They were rated on a 5-point Likert scale (1 = Strongly Disagree, 5 = Strongly Agree), with higher scores indicating stronger perceived research efficiency gains from generative AI. AI risk perception was measured using five items adapted from Q34_1 to Q34_5 of the Nature (2025) Global PhD Survey Questionnaire, covering concerns about AI weakening critical thinking and scientific writing skills, replacing core scientific competencies, displacing jobs in their fields, output quality and integrity, and output replicability and its implications for research. Items were rated on a 5-point Likert scale (1 = Not Worried at All, 5 = Extremely Worried), with higher scores indicating stronger risk perception. Item scores were averaged and then standardized.

Overtime work pressure was measured by three items, assessing the culture of long working hours, requirements for evening overtime, and work arrangements during weekends and public holidays. Items were rated on a 5-point Likert scale (1 = Strongly Disagree, 5 = Strongly Agree), with higher scores indicating greater perceived overtime work pressure. The scores were averaged and then standardized.

Control Variables

Control variables in this study covered two dimensions: demographic characteristics and educational background. Age and gender were included to control for individual attribute differences; year of doctoral study was included to reflect variations in research task intensity across different stages of training; doctoral degree type was included to account for potential differences in training orientation and career prospects; and geographic region of the respondents was included to mitigate potential biases arising from macro-level institutional environments.

Variable descriptions are presented in Table 1.

Table 1 Variable Measurement and Description

Scale	Item Code	Item Content Description		Source
Doctoral Research Satisfaction	7	Motivators	Independence and flexibility	Nature Q14a_2
	15		Guidance received about my research	Nature Q14a_5
	12		Ability to attend meetings and conferences	Nature Q14a_7
	8		Developing the skills I want to / need to	Nature Q14a_9

	9		The intellectual challenge	Nature Q14a_10
	13		Working with interesting and bright people	Nature Q14a_12
	10	Hygiene Factors	Social life and networks	Nature Q14a_1
	11		Overall relationship with supervisor	Nature Q14a_3
	14		Overall compensation and benefits	Nature Q14a_4
	16		Guidance received for my career	Nature Q14a_6
	17		Work-life balance	Nature Q14a_8
	18		Opportunities to travel / study overseas	Nature Q14a_11
Supportive Use	19	Experimental design		Nature Q33_4
	20	Literature tracking		Nature Q33_5
	21	Literature synthesis		Nature Q33_6
	22	Language polishing and refinement		Self-developed
Substitutive Use	23	Peer review		Nature Q33_1
	24	Article writing		Nature Q33_2
	25	Data collection and analysis		Nature Q33_3
	26	Job application preparation		Nature Q33_7
AI Efficiency Perception	27	Time cost savings		Self-developed
	28	Practical research benefits		Self-developed
	29	Subjective sense of research gain		Self-developed
AI Risk Perception	30	AI weakening critical thinking and scientific writing skills		Nature Q34_1
	31	AI replacing core scientific competencies		Nature Q34_2
	32	AI displacing jobs in their fields		Nature Q34_3
	33	Output quality and integrity		Nature Q34_4
	34	Output replicability and its implications for research		Nature Q34_5
Overtime Work Pressure	35	Culture of long working hours		Self-developed
	36	Requirements for evening overtime		Self-developed
	37	Work arrangements during weekends and public holidays		Self-developed
Control Variables	1	Age		Self-developed
	2	Gender		Self-developed
	3	Year of doctoral study		Self-developed
	4	Doctoral degree type		Self-developed
	6	Geographic region		Self-developed

Data

To ensure the quality of the items and the content validity of the questionnaire, expert review and pilot testing were conducted prior to the formal survey. First, three professors in educational psychology were invited to evaluate the initial items, assessing the match between each item and its corresponding construct, the reasonableness of the dimensional classification, and the clarity of the item wording. Based on the expert feedback, the wording of some items was revised. Second, 60 doctoral students were recruited for pilot testing to evaluate the comprehensibility and preliminary reliability of the questionnaire. In the pilot test, Cronbach's α coefficients for research satisfaction, generative AI use, AI efficiency perception, AI risk perception, and overtime work pressure were 0.872, 0.825, 0.797, 0.841, and 0.788, respectively. All reliability coefficients exceeded the acceptable threshold of 0.700, indicating

that the scales had acceptable internal consistency reliability. Based on the feedback from the pilot respondents, the final formal survey questionnaire was developed.

A questionnaire survey was administered to doctoral students currently enrolled at Chinese universities. Data were collected via Wenjuanxing from April to May 2026, with the questionnaire link distributed through doctoral student WeChat groups. The research purpose, anonymity principle, and data confidentiality commitment were explained on the front page of the questionnaire, and respondents participated voluntarily with informed consent. A total of 965 questionnaires were distributed, and 848 were returned. After excluding invalid responses (e.g., those with excessively short completion time, obvious patterned answering, or more than 10% missing values), 786 valid questionnaires were retained, yielding a valid response rate of 81.5%. The demographic characteristics of the sample are presented in Table 2.

Table 2. demographic characteristics

Variable	Category	Frequency	Percentage
Gender	Male	293	37.3%
	Female	493	62.7%
Age	18–24	143	18.2%
	25–34	546	69.5%
	35–44	76	9.7%
	45 above	21	2.7%
Year of study	1st year	248	31.6%
	2nd year	283	36.0%
	3rd year	201	25.6%
	4th year and above	54	6.9%
Discipline	STEM	335	42.6%
	Medical and health sciences	200	25.4%
	Humanities and social sciences	251	31.9%
Region	Eastern region	299	38.0%
	Central region	387	49.2%
	Western region	100	12.7%
Degree category	Academic doctorate	673	85.6%
	Professional doctorate	113	14.4%

Measurement Quality

Common method bias testing

Since all data were collected through self-report measures, common method bias might be a concern. Harman's single-factor test was conducted by entering all items of the core variables into an unrotated principal component analysis. The first factor accounted for 26.83% of the total variance, providing no indication of a dominant single-factor structure. Nevertheless, common method bias cannot be completely ruled out because the study relied primarily on self-reported measures.

Reliability testing

Cronbach's α coefficients were computed to assess the internal consistency reliability of the scales. As shown in Table 3, the Cronbach's α values for research satisfaction, generative AI use, AI efficiency perception, AI risk perception, and overtime work pressure were 0.912, 0.857, 0.824, 0.884, and 0.801, respectively. All coefficients exceeded the recommended threshold of 0.700,

indicating satisfactory internal consistency. Moreover, composite reliability (CR) values for all scales surpassed the 0.700 criterion, further confirming the reliability of the measures.

Table 3. Reliability test results for the scales

Scale	Number of items	Cronbach's α	CR
Research satisfaction	12	0.912	0.918
Generative AI use	8	0.857	0.862
AI Efficiency perception	3	0.824	0.836
AI Risk perception	5	0.884	0.891
Overtime work pressure	3	0.801	0.817

Validity testing

To examine the structural validity, convergent validity, and discriminant validity of the scales, a series of analyses were conducted. One half sample (n = 393) was used for exploratory principal component analysis (PCA), and the other half (n = 393) was used for confirmatory factor analysis (CFA).

First, Principal Component Analysis (PCA) was conducted to examine the dimensional structure of the scales. Four scales were unidimensional: research satisfaction, AI efficiency perception, AI risk perception, and overtime work pressure. Each scale yielded a single component, with KMO values ranging from 0.721 to 0.918, component loadings from 0.698 to 0.889, and cumulative variance explained from 55.51% to 73.42%. All Bartlett's tests were significant at $p < 0.001$ (Table 4).

Table 4. PCA Results for Unidimensional Scales

Scale	Item	Component loading	Eigenvalue	Cumulative variance explained
Research satisfaction	7	0. 750	6.661	55.51%
	8	0. 772		
	9	0. 786		
	10	0. 728		
	11	0. 756		
	12	0. 710		
	13	0. 776		
	14	0. 698		
	15	0. 734		
	16	0. 760		
	17	0. 746		
	18	0. 718		
AI efficiency perception	27	0. 869	2.103	70.09%
	28	0. 834		
	29	0. 797		
AI risk perception	30	0. 834	3.386	67.72%
	31	0. 816		
	32	0. 789		
	33	0. 768		

	34	0. 700		
Overtime work pressure	35	0. 889	2.203	73.42%
	36	0. 854		
	37	0. 823		

Note: KMO and Bartlett's test for each scale: research satisfaction (KMO = 0.906; χ^2 = 1,438.27, df = 66, p < 0.001); AI efficiency perception (KMO = 0.742; χ^2 = 243.86, df = 3, p < 0.001); AI risk perception (KMO = 0.831; χ^2 = 629.41, df = 10, p < 0.001); overtime work pressure (KMO = 0.708; χ^2 = 196.34, df = 3, p < 0.001).

The eight items of generative AI use were subjected to PCA. The KMO value was 0.853, and Bartlett's test of sphericity reached statistical significance (χ^2 = 930.72, df = 28, p < 0.001). Based on the eigenvalue-greater-than-one criterion, two components were identified, corresponding to supportive use and substitutive use, which together accounted for 69.38% of the total variance. The component loadings ranged from 0.729 to 0.854 (Table 5).

Table 5. PCA Results for Generative AI Use

Item	Component 1	Component 2
19	0.854	0.208
20	0.817	0.182
21	0.776	0.235
22	0.738	0.169
23	0.190	0.854
24	0.221	0.808
25	0.164	0.768
26	0.199	0.729

Second, CFA was conducted to evaluate the model fit of each scale, with $\chi^2/df < 3$, CFI > 0.90, TLI > 0.90, RMSEA < 0.08, and SRMR < 0.08 adopted as the criteria for good model fit. As shown in Table 6, all fit indices met the above criteria, indicating good structural validity.

Table 6. CFA Model Fit Indices

Scale	Model type	χ^2/df	CFI	TLI	RMSEA	SRMR
Research satisfaction	One-factor model	1. 842	0. 946	0. 934	0. 046	0. 038
Generative AI use	Two-factor model	1. 715	0. 958	0. 945	0. 043	0. 034
AI risk perception	One-factor model	1. 486	0. 978	0. 967	0. 035	0. 028
AI efficiency perception	One-factor model	1. 372	0. 989	0. 978	0. 031	0. 022
Overtime work pressure	One-factor model	1. 438	0. 982	0. 973	0. 033	0. 025

Third, convergent validity of the latent constructs was examined based on the CFA results. According to the criteria proposed by Fornell and Larcker, convergent validity is adequate when CR exceeds 0.700 and AVE exceeds 0.500. As shown in Table 7, all latent constructs had AVE values above 0.500 and CR values above 0.700, fully meeting the criteria for convergent validity.

Table 7. Convergent Validity Results for Latent Constructs

	AVE	CR
Research satisfaction	0.512	0.908
Generative AI use	0.503	0.852
AI risk perception	0.614	0.827

AI efficiency perception	0.568	0.883
Overtime work pressure	0.582	0.807

Fourth, discriminant validity was assessed to examine the conceptual distinctiveness among the latent constructs. The Fornell-Larcker criterion was first applied, which requires that the square root of the AVE for each latent construct exceeds its correlations with other constructs. As shown in Table 8, the square roots of the AVE for all latent constructs were greater than the absolute values of their inter-construct correlations, and all correlation coefficients were below 0.850. In addition, the heterotrait-monotrait ratio (HTMT) was employed as a supplementary method, and all HTMT values were below 0.850, further supporting discriminant validity.

Table 8. Latent Variable Correlation Matrix and Discriminant Validity Test

	Research satisfaction	Generative AI use	AI efficiency perception	AI risk perception	Overtime Work pressure
Research satisfaction	(0.716)				
Generative AI use	0.298***	(0.709)			
AI efficiency perception	0.331***	0.369***	(0.754)		
AI risk perception	−0.251***	0.216***	−0.172***	(0.784)	
Overtime work pressure	−0.184***	0.163***	−0.105*	0.312***	(0.763)

Model

This section explained how doctoral research satisfaction was estimated using the graded response model from item response theory, and then outlined the analytical strategy, which consisted of three layers. The direct nonlinear associations between generative AI use and research satisfaction were tested by examining the quadratic terms of supportive and substitutive use (H1 and H2). In addition, whether efficiency perception and risk perception showed conditional indirect effects in these associations was examined by testing the conditional indirect effects through these two mediators (H3 and H4). Furthermore, whether overtime work pressure moderated the quadratic associations between generative AI use and the two mediating variables was investigated by testing the quadratic interaction terms (H5).

Estimation of Research Satisfaction Latent Score

The graded response model (GRM) from item response theory was utilized to estimate doctoral students' latent research satisfaction scores. The GRM is a standard model for analyzing ordinal polytomous items such as Likert-type scales. As reported in the validity analysis (Table 4), research satisfaction exhibited a unidimensional structure, with one component extracted and a cumulative variance explained of 54.32%. This supported the assumption of unidimensionality required for the GRM.

Let Z_{ij} denote the observed score of individual i on item j (j=1,2,…..12), taking values k=1, 2, ..., K_j. The GRM characterized the relationship between the latent trait and ordinal responses by modeling cumulative probabilities. The probability that individual i scores at or above category k on item was given by:

$$P\left(Z_{ij} \geq k \mid \theta_i\right) = \frac{1}{1+exp[-a_j(\theta_i - b_{jk})]} \quad (1)$$

Where θ_i denoted the latent research satisfaction score; a_j was the discrimination parameter, reflecting the item's ability to distinguish individuals with different satisfaction levels; and b_{jk} was

the threshold parameter, reflecting the satisfaction level required for a response to shift from below category k to category k or higher. The probability of selecting a specific category k on an item j was the difference between adjacent cumulative probabilities: parameters were estimated using marginal maximum likelihood estimation, and individual latent scores were obtained using the expected a posteriori (EAP) method. The resulting latent satisfaction scores were then standardized (M=0, SD=1) to ensure comparability of scale and interpretability of results.

Main Effect Testing

To test H1 and H2, hierarchical regression analysis was employed to estimate the associations of supportive use and substitutive use with research satisfaction. Prior to model estimation, all predictor variables were standardized, and quadratic terms were constructed as the product of the standardized predictors. Because standardization inherently mean-centers the variables, no additional mean centering was performed. Descriptive statistics for the dependent and independent variables were reported in Table 9.

Table 9. Descriptive Statistics for Key Variables

Variable	Mean	SD	Min	Max
Research Satisfaction	0.000	1.000	−2.318	2.146
Supportive Use	0.000	1.000	−2.205	1.874
Supportive Use (squared)	0.978	1.126	0.000	5.412
Substitutive Use	0.000	1.000	−1.986	1.692
Substitutive Use (squared)	0.934	1.048	0.000	4.286

Supportive use and substitutive use were moderately correlated in Table 10.

Table 10. Pairwise Correlations among Key Variables

Variable	Research Satisfaction	Supportive Use	Supportive Use (squared)	Substitutive Use	Substitutive Use (squared)
Research Satisfaction	1.000				
Supportive Use	0.376	1.000			
Substitutive Use	0.184	0.412	1.000		
Supportive Use (squared)	−0.142	0.027	−0.063	1.000	
Substitutive Use (squared)	−0.081	−0.044	0.016	0.187	1.000

Variance inflation factors (VIFs) computed from the full joint regression model were all below 3, indicating that multicollinearity did not represent a substantive concern. The model specification consisted of three hierarchical levels. In the first level, control variables were entered. In the second level, the linear terms of the independent variables were added. In the third level, the squared terms of the independent variables were included. To enable direct comparison between the two types of use, a joint model was specified that simultaneously includes the linear and squared terms of both supportive and substitutive use. The specific forms were specified as follows:

Level 1：$F_i = \alpha_0 + \delta X_i + \varepsilon_i$

Level 2：$F_i = \alpha_1 + \beta_1 S_i + \delta X_i + \varepsilon_i$ （2）

Level 3：$F_i = \alpha_2 + \beta_1 S_i + \beta_2 S_i^2 + \delta X_i + \varepsilon_i$

Joint model: $F_i = \alpha_3 + \gamma_1 Supportive_i + \gamma_2 Supportive_i^2 + \gamma_3 Substitutive_i + \gamma_4 Substitutive_i^2 + \delta X_i + \varepsilon_i$

Where F_i denoted the composite score of research satisfaction, S_i denoted the score of either supportive use or substitutive use, X_i represented the vector of control variables, and ε_i was the random error term. For each predictor estimated separately, if β_1 was positive and statistically significant, and β_2 was negative and statistically significant, then the necessary conditions for an inverted U-shaped relationship were satisfied. In the joint model, $Supportive_i$ and $Substitutive_i$ were entered simultaneously, allowing the coefficients for the two types of use to be compared within the same model. The turning points for supportive and substitutive use were computed separately using the corresponding coefficients: γ_1 and γ_2 for supportive use, and γ_3 and γ_4 for substitutive use.

However, these conditions are necessary but not sufficient for an inverted U-shaped relationship (Lind & Mehlum, 2010). The Lind-Mehlum test based on the Sasabuchi composite hypothesis framework was therefore employed. This test examined whether the slope at the lower bound of the observed range was significantly positive and whether the slope at the upper bound was significantly negative. In addition, Fieller 95% confidence intervals were computed for each turning point, and the number of observations on each side of the turning point was calculated. As a sensitivity check, the quadratic regression models were re-estimated after excluding the upper 5% tail observations of each independent variable.

Mediation Effect Testing

H3 and H4 concerned the conditional indirect effects of AI efficiency perception and AI risk perception in the relationship between generative AI use and doctoral research satisfaction. To test these hypotheses, the Bootstrap method was employed for nonlinear mediation analysis, with 5,000 resampling iterations to estimate the instantaneous mediation effects. The testing procedure consisted of two steps. In the first step, the associations of supportive use and substitutive use with the two potential mediators were examined. In the second step, the mediating variables were respectively incorporated into the main effect models for supportive use and substitutive use. The models were specified as follows:

$$M_i = \alpha_3 + \theta_1 S_i + \theta_2 S_i^2 + \delta X_i + \varepsilon_i$$

$$F_i = \alpha_4 + \beta_1 S_i + \beta_2 S_i^2 + \lambda M_i + \delta X_i + \varepsilon_i \quad (3)$$

Where S_i denoted the score of either supportive use or substitutive use, and M_i represented either AI efficiency perception or AI risk perception. The Bootstrap method was used to estimate the 95% confidence intervals for the indirect effects. If a confidence interval did not contain zero, the conditional instantaneous indirect effect was considered statistically significant. Given the nonlinear relationship between the independent variable and the mediating variables, three representative values of the independent variable were selected (the mean minus one standard deviation, the mean, and the mean plus one standard deviation), and the instantaneous mediation effects were reported accordingly to illustrate the trend of the mediation effect as the independent variable varied.

Moderation Effect Testing

H5 concerned the positive moderating effect of overtime work pressure on the mediating pathways. To test this hypothesis, a hierarchical regression model with interaction terms was constructed. The moderating variable was standardized prior to model estimation, as were all other predictors. The models were specified as follows:

$$M_i = \alpha_5 + \theta_1 S_i + \theta_2 S_i^2 + \eta P_i + \rho_1 (S_i \times P_i) + \delta X_i + \varepsilon_i$$

$$M_i = \alpha_6 + \theta_1 S_i + \theta_2 S_i^2 + \eta P_i + \rho_2 (S_i^2 \times P_i) + \delta X_i + \varepsilon_i \quad (4)$$

Where P_i denoted the score of overtime work pressure, and $S_i \times P_i$ and $S_i^2 \times P_i$ represented the interaction terms between the moderating variable and the linear term and the squared term of the independent variable, respectively. If ρ_1 was statistically significant, it indicated that overtime work pressure moderated the linear relationship between generative AI use and the mediating variables. If ρ_2 was statistically significant, it indicated that overtime work pressure moderated the nonlinear relationship between generative AI use and the mediating variables. If neither was significant, the moderating effect was not supported.

Empirical Results

Main Effect Testing

To test the nonlinear associations of supportive use and substitutive use with doctoral research satisfaction, hierarchical regression analysis was employed. All regression models were estimated with robust standard errors. Independent variables were standardized and entered into the regression models sequentially. In Table 11, Model 1 included only the control variables. Models 2 and 3 successively added the linear term and the squared term of supportive use on the basis of the control variables. Models 4 and 5 did the same for substitutive use. The results indicated that the coefficients of the linear terms for both supportive and substitutive use were positive and statistically significant, while the coefficients of the squared terms were negative and statistically significant. These patterns satisfied only the necessary conditions for an inverted U-shaped relationship. However, separate model estimations cannot support valid comparisons between the two correlated constructs. Therefore, Model 6 was adopted as a joint model that simultaneously includes linear and squared terms of both constructs for rigorous cross-construct comparison. As shown in Table 11, both supportive and substitutive use exhibited significant inverted U-shaped associations with doctoral research satisfaction. The positive linear coefficients and negative squared coefficients remained statistically significant for both constructs even after mutually controlling for the alternative dimension of AI use.

Table 11. Effects of Generative AI Use on Doctoral Research Satisfaction

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Supportive Use		0.224*** （0.021）	0.219*** （0.022）			0.267*** (0.033)
Supportive Use Squared			-0.112** （0.036）			−0.145** (0.051)
Substitutive Use				0.138*** （0.020）	0.134*** （0.020）	0.100*** (0.022)
Substitutive Use Squared					-0.101** （0.034）	−0.082** (0.036)
Control Variables	YES	YES	YES	YES	YES	YES
R^2	0.124	0.161	0.169	0.142	0.149	0.195
ΔR^2		0.037***	0.008**	0.018***	0.007***	
F	65.42***	79.38***	75.26***	69.74***	68.15***	81.32***
Constant	0.031**	0.034**	0.038***	0.039***	0.043***	0.032***

	（0.011）	（0.011）	（0.010）	（0.010）	（0.009）	(0.009)

Note: Standard errors are in parentheses. *p < 0.05, **p < 0.01, ***p < 0.001. The same applies to the following tables.

To formally test whether the relationships followed a true inverted U-shaped pattern, the Lind--Mehlum test was conducted based on separate quadratic models to individually verify the existence of inverted U-shaped associations for each construct. As shown in Table 12, the test was statistically significant for both supportive use and substitutive use. For both predictors, the slope at the lower bound was significantly positive and the slope at the upper bound was significantly negative, satisfying the formal criteria for an inverted U-shaped relationship.

Cross-construct comparisons of turning points and their theoretical interpretations rely exclusively on the joint model, which accounts for the correlation between the two dimensions of generative AI use. The turning points estimated from the joint model were 0.92 for supportive use (Fieller 95% CI: [0.31, 1.21]) and 0.61for substitutive use (Fieller 95% CI: [0.20, 1.07]). Because the predictors were standardized, the turning points represent positions on the standardized scales. Converted back to the original composite metric, the turning point was 3.99 on the 1-5 scale for supportive use (M = 2.961, SD = 1.120) and 3.33 for substitutive use (M = 2.680, SD = 1.066). These values indicate the frequency of AI use at which the positive association with research satisfaction reaches its estimated maximum. Both turning points fell within the observed ranges, with observations distributed on both sides of the turning point (supportive: 76.8% below, 23.2% above; substitutive: 69.5% below, 30.5% above). The positive association with research satisfaction reached its maximum at a higher frequency of supportive use than of substitutive use.

To assess whether the declining segments were driven by a small number of upper-tail observations, additional robustness checks were conducted on the joint model. Although the upper bound for supportive use (1.874) was relatively close to its turning point (0.92), 23.2% of observations still fell above the turning point. Notably, the magnitude of the slope at the upper observed bound for supportive use was small. Even though the upper-bound slope was statistically negative within the observed data range, inferences for extremely high frequencies of supportive use should be interpreted with caution. The joint model was re-estimated after excluding the upper 5% and 10% tails of each predictor. The linear terms remained positive and significant, the squared terms remained negative and significant, and the turning points were close to joint-model estimates. These results indicated that the negative slopes beyond the turning points were not driven solely by a small number of extreme upper-tail observations. Thus, H1 and H2 were supported.

Table 12. Lind-Mehlum Inverted-U test and Turning-point Statistics

	Supportive AI Use	Substitutive AI Use	Joint Model
Lower Bound	−2.205	−1.986	−2.205
Slope at Lower Bound	0.906*	0.426***	0.872*
Lower Bound Slope 95% CI	[0.184, 1.628]	[0.168, 0.684]	[0.156, 1.588]
Upper Bound	1.874	1.692	1.874
Slope at Upper Bound	−0.276*	−0.177**	−0.258*
Upper Bound Slope 95% CI	[−0.512, −0.040]	[−0.328, −0.026]	[−0.486, −0.030]
Lind–Mehlum Test	0.019	0.007	Supportive: 0.022 Substitutive: 0.009

Turning Point	0.92	0.61	Supportive: 0.92 Substitutive: 0.61
Fieller 95% CI	[0.31, 1.21]	[0.20, 1.07]	Supportive: [0.31, 1.21] Substitutive: [0.20, 1.07]
Below Turning Point	604 (76.8%)	546 (69.5%)	Supportive: 604 (76.8%) Substitutive: 546 (69.5%)
Above Turning Point	182 (23.2%)	240 (30.5%)	Supportive: 182 (23.2%) Substitutive: 240 (30.5%)

The above results were consistent with the expectations of cognitive load theory and further revealed differences in the effect boundaries between the two usage modes. In supportive use, AI mainly undertook auxiliary cognitive operations while doctoral students retained control over the research process, and therefore the erosion of cognitive capacity remained relatively limited (Kim et al., 2025). By contrast, substitutive use involved deep intervention of AI in core research activities, and doctoral students lost opportunities to accumulate academic competencies through hands-on practice. As a result, its inverted U-shaped curve reached the peak earlier, and marginal benefits turned negative more quickly.

Mediation Effect Testing

To examine the conditional instantaneous indirect effects of AI efficiency perception and AI risk perception in the relationship between generative AI use and doctoral research satisfaction, a combination of hierarchical regression and Bootstrap methods was adopted. Table 13 presented the hierarchical regression results. Models 7 through 10 showed that both supportive use and substitutive use had significant inverted-U-shaped associations with AI efficiency perception, and AI efficiency perception in turn had a significant positive association with research satisfaction. Models 11 through 14 showed that both types of generative AI use had a positive and predominantly linear association with AI risk perception, with significant positive squared terms, and AI risk perception in turn had a significant negative association with research satisfaction. These hierarchical-regression outcomes satisfied the prerequisites for mediation analysis.

Table 13. Mediation Effects of Generative AI Use on Research Satisfaction

Variable	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12	Model 13	Model 14
Supportive Use	0.192*** （0.022）	0.174*** （0.023）			0.075*** (0.018)	0.070*** （0.018）		
Supportive Use Squared	-0.078** （0.032）	-0.088** （0.036）			0.052*** （0.021）	0.052*** (0.021)		
Substitutive Use			0.114*** （0.020）	0.096*** （0.021）			0.098*** (0.020)	0.092*** (0.021)
Substitutive Use Squared			-0.058* （0.029）	-0.074** （0.030）			0.080** (0.024)	0.080** (0.024)
AI Efficiency Perception		0.218*** （0.033）		0.168*** （0.032）				
AI Risk Perception						-0.161*** （0.031）		-0.158*** （0.032）

Control Variables	YES	YES	YES	YES	YES	YES	YES	YES
R^2	0.108	0.226	0.090	0.191	0.056	0.229	0.075	0.197
F	44.18***	87.65***	37.24***	74.36***	29.06***	88.47***	32.84***	76.05***
Constant	0.024*** （0.006）	0.031*** （0.009）	0.019** （0.006）	0.037*** （0.008）	0.019** （0.007）	0.033*** （0.006）	0.022*** （0.008）	0.039*** （0.009）

Given the relatively limited statistical power of the traditional stepwise approach, the Bootstrap method with 5,000 resampling iterations was employed to estimate conditional instantaneous indirect effects for the quadratic-specification associative-pathway model. In this quadratic-specification model, conditional instantaneous indirect effects were computed as$(\theta_1 + 2\theta_2 S) \times \lambda$, where θ_1 and θ_2 represented the linear and quadratic coefficients for predicting the mediator, and λ denoted the coefficient of the mediator on the outcome variable. This derivative-based conditional instantaneous indirect effect was computed within each bootstrap replicate, and percentile-based 95% bootstrap confidence intervals were constructed from the resulting bootstrap distribution.

Because the relationship between the independent variables and the mediating variables was nonlinear, the magnitude of the conditional instantaneous indirect effects varied across different frequencies of AI use. Conditional instantaneous indirect effects were therefore estimated at three representative values: the mean minus one standard deviation, the mean, and the mean plus one standard deviation. Table 14 presents the results. For AI efficiency perception, the conditional instantaneous indirect effect was positive and significant at M−1SD and M, and became non-significant at M+1SD, with the magnitude decreasing as the frequency of AI use increased. For AI risk perception, the conditional instantaneous indirect effect became increasingly negative as the frequency of AI use increased, reaching significance at M and M+1SD. These results supported H3 and H4: AI efficiency perception showed a positive conditional instantaneous indirect effect that weakened as the frequency of AI use increased, whereas AI risk perception showed a negative conditional instantaneous indirect effect that strengthened as the frequency of AI use increased.

Table 14. Bootstrap Results for Conditional Instantaneous Indirect Effects

Independent Variable	Mediating Variable	M−1SD [95% CI]	M [95% CI]	M+1SD [95% CI]
Supportive Use	AI Efficiency Perception	0.076*** [0.036, 0.118]	0.042*** [0.018, 0.068]	0.008 [−0.001, 0.017]
	AI Risk Perception	0.005 [−0.003, 0.013]	−0.012* [−0.023, −0.002]	−0.029*** [−0.048, −0.012]
Substitutive Use	AI Efficiency Perception	0.039*** [0.017, 0.063]	0.019* [0.003, 0.036]	−0.000 [−0.011, 0.011]
	AI Risk Perception	0.010 [−0.004, 0.024]	−0.016* [−0.030, −0.003]	−0.041*** [−0.070, −0.019]

Moderation Effect Testing

To test the positive moderating effect of overtime work pressure on the mediating pathways, interaction terms were incorporated into the regression models, and the moderating variable was standardized prior to model estimation, consistent with the treatment of all other predictors. In Table 15, the interaction terms between the squared terms of both supportive use and substitutive use and

overtime work pressure were significantly positive across Models 15–18, whereas the interaction terms involving the linear terms were not significant. These results indicated that overtime work pressure positively moderated the relationship between generative AI use and both mediating variables, with the moderating effect operating on the quadratic terms rather than the linear terms. Specifically, AI efficiency perception exhibited an inverted-U-shaped association with generative AI use given its negative squared coefficient. Under high overtime work pressure, this inverted-U curve was altered and its turning point shifted to the right. By contrast, AI risk perception yielded a positive squared coefficient, indicating a U-shaped association with generative AI use. Under high overtime work pressure, this U-shaped association became more pronounced, such that risk perception grew at an accelerating pace as the frequency of generative AI use increased. Thus, Hypothesis H5 was supported.

Table 15. Moderation Effects of Overtime Work Pressure

Variable	AI Efficiency Perception		AI Risk Perception	
	Model 15	Model 16	Model 17	Model 18
Supportive Use	0.184*** （0.023）		-0.052** （0.019）	
Supportive Use Squared	-0.082** （0.032）		0.064** （0.025）	
Substitutive Use		0.108*** （0.021）		0.136*** （0.021）
Substitutive Use Squared		-0.056* （0.029）		0.088*** （0.027）
Overtime Work Pressure	0.131*** （0.020）	0.114*** （0.019）	0.091*** （0.018）	0.083*** （0.017）
Supportive Use × Overtime Work Pressure	0.033 （0.029）		0.027 （0.026）	
Supportive Use Squared × Overtime Work Pressure	0.085** （0.037）		0.078** （0.035）	
Substitutive Use × Overtime Work Pressure		0.031 （0.027）		0.025 （0.026）
Substitutive Use Squared × Overtime Work Pressure		0.083** （0.034）		0.075** （0.032）
Control Variables	YES	YES	YES	YES
R^2	0.172	0.145	0.101	0.116
F	60.18***	48.73***	33.45***	37.28***
Constant	0.026*** （0.006）	0.023*** （0.007）	0.019** （0.006）	0.024*** （0.006）

To further explain how overtime work pressure shaped the inverted-U relationship for AI efficiency perception, turning-point movement for this specific pathway was calculated under different levels of overtime work pressure following the approach of Haans et al. (2016).Taking the moderating pathway of supportive use on AI efficiency perception as an example (Model 15), the moderated function was differentiated with respect to supportive use (x), and the derivative was set

to zero to obtain the turning-point expression. Further differentiating the turning point with respect to overtime work pressure (W) yields $dx*/dW = 0.0184 / (-0.082 + 0.085W)^2 > 0$, indicating that the turning point shifted progressively to the right as overtime work pressure increased. This implied that under high overtime work pressure, doctoral students required a higher frequency of generative AI use to reach the peak value of AI efficiency perception.

Robustness Checks

To ensure the reliability of the main effect findings, robustness checks were performed from two aspects. First, the measurement of the dependent variable was altered. In the main effect analysis, research satisfaction was re-measured as the arithmetic mean of the 12 items. Models 19 and 20 in Table 16 presented the results after replacing the dependent variable measurement. The linear terms of both supportive use and substitutive use remained positively significant, and the squared terms remained negatively significant, consistent with the baseline results. Second, the sample range was adjusted. To assess whether the results were driven by extreme observations, the upper 5% tails of each predictor were excluded. Models 21 and 22 showed that, after the sample adjustment, the linear terms of both supportive use and substitutive use remained positively significant, and the squared terms remained negatively significant. These results were consistent with the baseline regressions, indicating that the findings were robust.

Table 16. Robustness Checks

Variable	Model 19	Model 20	Model 21	Model 22
	Alternative Dependent Variable Measurement		Adjusted Sample Range	
Supportive Use	0.202*** （0.027）		0.207*** （0.028）	
Supportive Use Squared	-0.098** （0.038）		-0.101** （0.039）	
Substitutive Use		0.121*** （0.025）		0.125*** （0.026）
Substitutive Use Squared		-0.085** （0.033）		-0.089** （0.034）
Control Variables	YES	YES	YES	YES
R^2	0.158	0.137	0.157	0.142
F	22.74***	20.41***	23.16***	20.83***
Constant	0.031*** （0.009）	0.036*** （0.008）	0.033*** （0.008）	0.037*** （0.008）
N	786	786	746	746

Heterogeneity Analysis

To examine whether the main findings varied across different disciplinary groups, the full sample was divided into three groups according to doctoral students' research fields: STEM disciplines, medical and health disciplines, and humanities and social sciences. Subsample regression analyses were first conducted separately for the three groups. As shown in Models 23–28 in Table 17, the linear terms of both supportive and substitutive use were significantly positive, while the squared terms were significantly negative across all three disciplinary contexts. These results demonstrated that the inverted U-shaped associations between generative AI use and doctoral

research satisfaction were robust across the three disciplinary groups, consistent with the baseline findings. In terms of effect magnitude, the estimated peak of doctoral research satisfaction at the turning point for supportive use was highest in STEM disciplines (0.101), followed by humanities and social sciences (0.089) and medical and health disciplines (0.080). A similar pattern was observed for substitutive use, with estimated peaks of 0.041 in STEM, 0.038 in humanities and social sciences, and 0.028 in medical and health disciplines.

Table 17. Heterogeneity Analysis by Disciplinary Background

Variable	Model 23	Model 24	Model 25	Model 26	Model 27	Model 28
	STEM Disciplines		Medical and Health Disciplines		Humanities and Social Sciences	
Supportive Use	0.213*** （0.034）		0.167*** （0.046）		0.185*** （0.041）	
Supportive Use Squared	-0.112** （0.043）		-0.087** （0.042）		-0.096** （0.040）	
Substitutive Use		0.124*** （0.019）		0.098*** （0.038）		0.114*** （0.036）
Substitutive Use Squared		-0.095** （0.035）		-0.086** （0.036）		-0.086** （0.034）
Control Variables	YES	YES	YES	YES	YES	YES
R^2	0.172	0.143	0.131	0.113	0.148	0.132
F	14.32***	12.45***	7.26***	6.14***	10.58***	9.33***
Sample Size	335	335	335	335	251	251

The descriptive variation in the estimated peaks can be attributed to the nature of research practices across fields. STEM research involves a substantial volume of repetitive tasks such as data analysis, coding, and experimental simulations. In this context, generative AI, functioning as a cognitive collaborator, effectively alleviates the burden of mechanical research work, yielding pronounced gains in cognitive resource release. However, when AI substitutes for core analytical processes, it also encroaches upon doctoral students' opportunities for hands-on practice, resulting in a steeper decline on the descending side of the inverted U-shaped curve. By contrast, in the humanities and social sciences, doctoral students predominantly employ generative AI for auxiliary tasks such as literature organization, text refinement, and perspective synthesis, rather than for core analytical or inferential work. As these tasks do not substantially intrude upon the essential processes of critical thinking and theoretical reasoning, the risk of cognitive displacement remains relatively contained. Consequently, the peak effect in this group falls between that of STEM and medical and health sciences. The pattern is different again in medical and health sciences, where research relies heavily on clinical observation and experiential accumulation, and the scope for AI integration remains relatively limited. As a result, both the gains and risks associated with supportive and substitutive AI use are more moderate in this disciplinary context.

Conclusion, Limitation and Implication

Conclusion

Drawing on cognitive load theory and self-determination theory, and using data from a questionnaire survey of Chinese doctoral students, this study examined how two dimensions of

generative AI use were associated with doctoral research satisfaction, and how AI efficiency perception, AI risk perception, and overtime work pressure were involved in these associations. The main findings are as follows.

Both supportive use and substitutive use showed significant inverted U-shaped associations with doctoral research satisfaction: moderate frequencies of use were associated with higher satisfaction, whereas excessively high frequencies were associated with lower satisfaction. In addition, the turning point for supportive use was higher than that for substitutive use, indicating that the positive association with research satisfaction reached its maximum at a higher frequency of supportive use, whereas substitutive use reached its peak earlier. Heterogeneity analysis indicated that the inverted U-shaped association held across STEM, medical and health, and humanities and social sciences disciplines. For both types of use, the estimated peak at the turning point was highest in STEM disciplines, followed by humanities and social sciences, and lowest in medical and health disciplines. Despite this variation, the pattern of association was similar across the three disciplinary contexts. This finding extends prior research that has predominantly framed generative AI use in terms of either benefits or risks. The association between generative AI use and doctoral research satisfaction varies with both the frequency of use and the type of use orientation. Accordingly, studies on how generative AI shapes doctoral students' research experiences should distinguish among different modes of AI use instead of treating AI use as a single and homogeneous behavior.

The association between AI use and doctoral research satisfaction operated differently through the two pathways. AI efficiency perception showed a positive conditional instantaneous indirect effect, whereas AI risk perception showed a negative one. The two pathways exhibited different patterns across different frequencies of AI use. The result indicates that AI use is not only a matter of cognitive resource reallocation but also involves students' subjective experience of competence and autonomy. It helps explain why the two pathways diverge across different frequencies of AI use and provides support for the integrated framework of cognitive load theory and self-determination theory.

Moderation analysis indicated that overtime work pressure positively moderated the quadratic associations between AI use and both AI efficiency perception and AI risk perception, with the moderating effect operating on the quadratic terms rather than the linear terms. That is to say, the psychological effects of generative AI use are embedded in the specific contexts of doctoral students' research work. AI is not an isolated technological input, and its perceived value and risks may vary with the level of work pressure doctoral students experience. Understanding the role of generative AI use in doctoral education requires attention to the research working conditions.

Limitation

Several limitations should be acknowledged.

Research design and causal inference. Given the cross-sectional nature of the data, the reported relationships should be interpreted as statistical associations rather than causal effects. Accordingly, reverse causality cannot be ruled out: more satisfied doctoral students may be more willing to use AI, whereas less satisfied students may rely on it more heavily to cope with research difficulties. In addition, unobserved factors such as supervisory relationships, the institutional research environment, and individual research competence may jointly affect AI use and research satisfaction.

Future research could employ longitudinal or experimental designs with more comprehensive controls to examine temporal ordering and causal mechanisms more rigorously.

Measurement error and common method bias. AI efficiency perception, AI risk perception, overtime work pressure, and research satisfaction were all measured in the same questionnaire, and shared response tendencies or subjective evaluation bias may have influenced the observed associations. In addition, the survey relied on self-reported AI use frequency, which may not fully capture actual usage behavior. Future research could use multi-source data, objective behavioral indicators, or temporally separated measures to reduce common method bias.

Sampling design and generalizability. The sample was drawn from doctoral students currently enrolled at Chinese universities, which may limit the generalizability of the findings to doctoral populations in other cultural and institutional contexts. In addition, the disciplinary distribution was uneven, with STEM students overrepresented relative to medical and health and humanities and social sciences students, which may constrain the robustness of the disciplinary comparisons. Item interpretations may also vary across disciplines, and AI perceptions and use may differ by institutional environment, digital infrastructure, and disciplinary norms. Future research should adopt more representative and balanced samples and examine whether the findings hold across different national and institutional contexts.

Implication

The findings of this study offer several practical implications for doctoral education and institutional support in the age of generative AI. These suggestions should be viewed as potential directions for practice rather than as empirically validated interventions.

First, guidance and training in the use of generative AI. Both supportive and substitutive use showed inverted U-shaped associations with doctoral research satisfaction. Moderate frequencies were associated with higher satisfaction, whereas excessive frequencies were associated with lower satisfaction. Based on this finding, one potential direction for practice is to deepen doctoral students' understanding of both the benefits and limitations of generative AI and to integrate discussions of appropriate boundaries for AI-assisted research into research methods training. Universities may offer seminars, case studies, and hands-on workshops to help doctoral students develop informed judgments about AI use (Ma et al., 2024). Supervisors may also discuss reasonable boundaries of AI use during routine research guidance, exploring how generative AI can function more as a cognitive amplifier than as a substitute for core research activities.

Second, academic ethics education and curriculum development. Higher AI use frequency was associated with stronger AI risk perception, which in turn was negatively associated with research satisfaction. Accordingly, one direction worth further consideration is to update courses on academic ethics and research integrity to address ethical issues related to generative AI use. These may include standards for citing and disclosing AI-generated content, as well as quality verification and accountability for AI-generated outputs (Tian & Xiao, 2024). Such curriculum design can help doctoral students develop critical judgment about AI use in specific research contexts and consider the balance between technological assistance and academic integrity.

Third, research evaluation and management. Overtime work pressure positively moderated the quadratic associations between AI use and both AI efficiency perception and AI risk perception. Based on this moderating result, one possible direction for institutional support is to further attend

to doctoral students' work schedules and task demands, and to examine whether research-related pressure is associated with the frequency and mode of their AI use. In research evaluation, one direction worth exploring is the use of formative assessment indicators. These could focus on research design rigor, methodological appropriateness, and argumentation completeness (Yu & Liu, 2025), offering a testable way to reduce reliance on publication counts alone.

Data Availability Statement

The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

References

Alireza, A., Yang, J. Q., & Syed, M. F. (2026). Knowledge workers' trust and reception of generative AI's advice in complex tasks. *International Journal of Information Management*, 88, 103031. https://doi.org/10.1016/j.ijinfomgt.2026.103031

Andersen, J. P., Degn, L., Fishberg, R., Graversen, E. K., Horbach, S. P. J. M., Kalpazidou Schmidt, E., Schneider, J. W., & Sørensen, M. P. (2025). Generative artificial intelligence (GenAI) in the research process: A survey of researchers' practices and perceptions. *Technology in Society*, 81, 102813. https://doi.org/10.1016/j.techsoc.2025.102813

Baber, H., Nair, K., Gupta, R., & Gurjar, K. (2024). The beginning of ChatGPT: A systematic and bibliometric review of the literature. *Information and Learning Sciences*, 125(8), 587-614. https://doi.org/10.1108/ILS-04-2023-0035

Chen, O., Paas, F., & Sweller, J. (2023). A cognitive load theory approach to defining and measuring task complexity through element interactivity. *Educational Psychology Review*, 35(2), 63-89. https://doi.org/10.1007/s10648-023-09782-w

Dai, Y., Lai, S., Lim, C. P., et al. (2023). ChatGPT and its impact on research supervision: Insights from Australian postgraduate research students. *Australasian Journal of Educational Technology*, 39(4), 74-88. https://doi.org/10.14742/ajet.8843

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11(4), 227-268. https://doi.org/10.1207/S15327965PLI1104_01

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., Stadler, M., Weller, J., Kuhn, J., & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Fauzi, F., Tuhuteru, L., Sampe, F., Ausat, A. M. A., & Hatta, H. R. (2023). Analyzing the role of ChatGPT in improving student productivity in higher education. *Journal on Education*, 5(4), 14886-14891. https://doi.org/10.31004/joe.v5i4.2563

Gao, Z., Cheah, J. H., Lim, X. J., & Luo, X. (2024). Enhancing academic performance of business students using generative AI: An interactive-constructive-active-passive (ICAP) self-determination perspective. *The International Journal of Management Education*, 22(2), 100958. https://doi.org/10.1016/j.ijme.2024.100958

Guleria, A., Krishan, K., Sharma, V., & Kanchanet, T. (2023). ChatGPT: Ethical concerns and challenges in academics and research. *The Journal of Infection in Developing Countries*, 17, 1292-1299. https://doi.org/10.3855/jidc.18738

Guo, S. H., & Zhu, D. L. (2025). Empowerment or crisis: The impact of generative AI use on postgraduate research capacity. *Tsinghua Journal of Education*, 46(6), 143-152. https://doi.org/10.14138/j.1001-4519.2025.06.014310

Haans, R. F. J., Pieters, C., & He, Z. L. (2016). Thinking about U: Theorizing and testing U- and inverted U-shaped relationships in strategy research. *Strategic Management Journal*, 37(7), 1177-1195. https://doi.org/10.1002/smj.2399

He, X., Huang, H., & Li, J. (2026). Depend less, think more: How AI knowledge empowers doctoral students' innovation behavior. *Education and Information Technologies*, 31, 2099-2122. https://doi.org/10.1007/s10639-025-13882-1

Herzberg, F., Mausner, B., & Snyderman, B. B. (1959). The motivation to work. John Wiley and Sons.

Hill, C., & Thabet, R. (2021). Publication challenges facing doctoral students: Perspective and analysis from the UAE. *Quality in Higher Education*, 27(3), 324-337. https://doi.org/10.1080/13538322.2020.1867036

Khalifa, M., & Albadawy, M. (2024). Using artificial intelligence in academic writing and research: An essential productivity tool. *Computer Methods and Programs in Biomedicine Update*, 5, 100145. https://doi.org/10.1016/j.cmpb.2024.100145

Khan, A., & Nasim, S. (2025). Mapping research on the subjective well-being of knowledge workers: A systematic enquiry deploying bibliometric. *Management Review Quarterly*, 75(1), 911-954. https://doi.org/10.1007/s11301-023-00399-5

Khan, O., & Ullah, F. (2024). Impact of ChatGPT on academic engagement and personalized learning among post-graduate students: A literature review. *Research Journal of Psychology*, 2(3), 153-179.

Kim, J., Yu, S., & Detrick, R. (2025). Exploring students' perspectives on generative AI-assisted academic writing. *Education and Information Technologies,* 30(1), 1265-1300. https://doi.org/10.1007/s10639-024-12878-7

Kirschner, P. A., Sweller, J., Kirschner, F., & Zambrano, R. J. (2018). From cognitive load theory to collaborative cognitive load theory. *International Journal of Computer-Supported Collaborative Learning*, 13, 213-233. https://doi.org/10.1007/s11412-018-9277-y

Li, Y., Sadiq, G., Qambar, G., & Zheng, P. Y. (2024). The impact of students' use of ChatGPT on their research skills: The mediating effects of autonomous motivation, engagement, and self-directed learning. *Education and Information Technologies*, 30, 1-32. https://doi.org/10.1007/s10639-024-12981-9

Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I., & Pechenkina, E. (2023). Generative AI and the future of education: Ragnarok or reformation? A paradoxical perspective from management educators. *The International Journal of Management Education*, 21(2), 100790. https://doi.org/10.1016/j.ijme.2023.100790

Lind, J. T., & Mehlum, H. (2010). With or without U? The appropriate test for a U-shaped relationship. *Oxford Bulletin of Economics and Statistics*, 72(1), 109-118. https://doi.org/10.1111/j.1468-0084.2009.00569.x

Lo, C. K., Hew, K. F., & Jong, M. S. Y. (2024). The influence of ChatGPT on student engagement: A systematic review and future research agenda. *Computers & Education*, 219, 105100. https://doi.org/10.1016/j.compedu.2024.105100

Ma, Y. Q., Huang, H., Wu, L., et al. (2024). "Technology empowerment" or "tool dependence": The impact of generative artificial intelligence on graduate students' research innovation capability. *e-Education Research*, 45(12), 58-66. https://doi.org/10.13811/j.cnki.eer.2024.12.008

Meng, N., Deli, M. M., & Rauf, U. A. A. (2025). Educational technology and AI: Bridging cognitive load and learner engagement for effective learning. *Sage Open*, 15(4), 1-21. https://doi.org/10.1177/21582440251395930

Mohammad, H., Maya, M., & David, B. R. (2026). Benefits and risks of using AI agents in research. *Hastings Center Report*, 56(1), 13-17. https://doi.org/10.1002/hast.70025

Pereira, R., Reis, I. W., Ulbricht, V., & dos Santos, N. (2024). Generative artificial intelligence and academic writing: An analysis of the perceptions of researchers in training. Management Research: *Journal of the Iberoamerican Academy of Management*, 22(4), 429-450. https://doi.org/10.1108/MRJIAM-01-2024-1501

Rahman, M. M., & Watanobe, Y. (2023). ChatGPT for education and research: Opportunities, threats, and strategies. *Applied Sciences*, 13, 5783. https://doi.org/10.3390/app13095783

Ray, P. P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems*, 3, 121-154. https://doi.org/10.1016/j.iotcps.2023.04.003

Rice, S., Crouse, S. R., Winter, S. R., & Rice, C. (2024). The advantages and limitations of using ChatGPT to enhance technological research. *Technology in Society*, 76, 102426. https://doi.org/10.1016/j.techsoc.2023.102426

Ringo, D. S. (2025). The effect of generative AI use on doctoral students' academic research progress: The moderating role of hedonic gratification. *Cogent Education*, 12(1), 2475268. https://doi.org/10.1080/2331186X.2025.2475268

Safdar, M., Siddique, N., Gulzar, A., Yasin, H., & Khan, M. A. (2024). Does ChatGPT generate fake results? Challenges in retrieving content through ChatGPT. *Digital Library Perspectives*, 40(4), 668-680. https://doi.org/10.1108/DLP-01-2024-0006

Stefano, B., Moritz, M., & Pierre, P. (2025). Drivers and barriers of AI adoption and use in scientific research. *Technological Forecasting and Social Change*, 220, 124303. https://doi.org/10.1016/j.techfore.2025.124303

Sverdlik, A., Hall, N. C., McAlpine, L., & Hubbard, K. (2018). The PhD experience: A review of the factors influencing doctoral students' completion, achievement, and well-being. *International Journal of Doctoral Studies*, 13, 361-388. https://doi.org/10.28945/4113

Tian, X. P., & Xiao, Z. Q. (2024). Academic ethics and risk prevention of generative AI-empowered graduate research writing. *Modern Educational Technology*, 34(8), 23-32. https://doi.org/10.3969/j.issn.1009-8097.2024.08.003

Uludağ, F., Kılıç, E., & Çelik, H. E. (2025). Artificial intelligence, social influence, and AI anxiety: Analyzing the intentions of science doctoral students to use ChatGPT with PLS-SEM. *Humanities and Social Sciences Communications*, 12(1), 1-17. https://doi.org/10.1057/s41599-025-05641-x

Van, T. T. P., Glorieux, A., Minnen, J., & Spruyt, B. (2024). Caught between academic calling and academic pressure? Working time characteristics, time pressure and time sovereignty predict PhD students' research engagement. *High Education*, 87, 1885-1904. https://doi.org/10.1007/s10734-023-01096-8

Yu, J. F., & Liu, H. Q. (2025). Exploring the integrated innovation of graduate education based on the new characteristics of generative artificial intelligence. *Academic Degrees & Graduate Education*, 9, 28-34. https://doi.org/10.16750/j.adge.2025.09.003

Zhou, J., Müller, H., Holzinger, A., & Chen, F. (2024). Ethical ChatGPT: Concerns, challenges, and commandments. *Electronics*, 13(17), 3417. https://doi.org/10.3390/electronics13173417

Zhou, S., & Zhang, S. (2024). Risks of generative AI application in academic research and their prevention mechanisms. *Heilongjiang Researches on Higher Education*, 42(3), 1-7. https://doi.org/10.3969/j.issn.1003-2614.2024.03.001

Zhu, H., & Yang, S. (2026). The impact of generative AI use on graduate students' research competence: The mediating role of critical thinking and the moderating role of research self-efficacy. *Behavioral Sciences,* 16(2), 304. https://doi.org/10.3390/bs16020304

Zou, M., & Huang, L. (2024). The impact of ChatGPT on L2 writing and expected responses: Voice from doctoral students. *Education and Information Technologies*, 29(11), 13201-13219. https://doi.org/10.1007/s10639-023-12397-x

Zou, Q. J., & Zhao, S. K. (2025). Exploring the impact of AI on research practices: A post digital ethnographic study of doctoral students' engagement with ChatGPT. *Studies in Continuing Education*, 1-17. https://doi.org/10.1080/0158037X.2025.2535565