Ethical risks and institutional responses of generative artificial intelligence empowering college teaching

LIU Honghao[1], XIA Pinyi[1], LI Yitong[2], HU Junhuan[2]

(1. Tianjin College, University of Science and Technology Beijing, Baodi 301830, Tianjin, China; 2. Pingliang College of Information Engineering, Pingliang 744000, Gansu, China; 3. Jiaying University, Meizhou 514015, Guangdong, China)

Abstract: Generative artificial intelligence is accelerating its embedding into the whole process of teaching in colleges and universities. While releasing efficiency dividends in the links of lesson preparation, teaching, evaluation and learning support, it also causes ethical risks such as the dissolution of academic integrity, the materialization of teacher-student relationship, the solidification of algorithm bias and the erosion of teaching autonomy. Based on the normative analysis framework of educational leadership ethics, this paper focuses on three dimensions: caring ethics, justice ethics and responsibility ethics, and constructs the analysis chain of ' risk identification-ethical attribution-institutional response '. This study adopts a speculative path combining normative analysis and policy text analysis to systematically sort out the typical ethical conflicts of generative artificial intelligence-enabled college teaching, and reveals the deep causes behind it, such as the imbalance of power structure, the fracture of responsibility chain and the absence of value rationality. On this basis, the system response strategy is proposed from four aspects : ethical regulation, organizational system, teaching governance and capacity building. It is advocated to replace " remedial ethics " with " preventive ethics " and reconstruct the teaching governance pattern of colleges and universities with " human-machine collaborative leadership. " The research aims to provide theoretical reference and institutional suggestions for colleges and universities to adhere to the education standard and realize the dynamic balance between technology empowerment and ethical protection in the era of artificial intelligence.

Keywords: generative artificial intelligence ; college teaching ; ethical risks ; education leadership ethics ; institutional response

DOI: https://doi.org/10.65196/

1. Introduction

1.1 Research background and questions raised

The generative artificial intelligence represented by ChatGPT and Wen Xin Yi Yan is accelerating its embedding into the whole teaching process of colleges and universities, from curriculum design, teaching resource development to formative evaluation and personalized learning support, and its application scenarios continue to expand. [1] The efficiency dividend released by technology empowerment is obvious to all: teachers can devote more energy to teaching design optimization and deep interaction between teachers and students, and learners can also obtain more immediate feedback support. However, technological empowerment and ethical risks always coexist. [2] The ambiguity of the attribution of the generated content impacts the foundation of academic integrity, the normalization of human-computer interaction dilutes the emotional connection between teachers and students, the deviation of training data solidifies the injustice of

Author Biographies
LIU Honghao (1998–), female, bachelor's degree, Research Assistant, specializing in Educational Administration.
XIA Pinyi (1995–), female, master's degree, Lecturer, specializing in Educational Administration.
LI Yitong (1997–), male, bachelor's degree, Lecturer, specializing in Educational Administration.
HU Junhuan (1998–), male, doctoral candidate, Teaching Assistant, specializing in Educational Administration.

teaching through algorithms, and the dependence of teachers on technical tools quietly erodes their professional judgment. The above risks are not isolated technical problems, but reflect the deep tension of educational leadership ethics in the intelligent era. Therefore, this paper puts forward the core research question : what ethical risks do generative artificial intelligence enable college teaching? How does educational leadership ethics provide a normative basis for institutional response?

1.2 Research purpose and significance

The purpose of this paper is to systematically identify the ethical risks of generative artificial intelligence empowering college teaching, reveal its ethical attribution, and derive institutional response strategies based on the normative analysis framework of educational leadership ethics.In the theoretical sense, this study introduces caring ethics, justice ethics and responsibility ethics into the context of intelligent technology, and expands the application boundary of educational leadership ethics. In the practical sense, the research provides institutional design reference for colleges and universities to formulate teaching ethics norms and reconstruct teaching governance pattern in the era of artificial intelligence.

1.3 Overview of research methods and ideas

This paper adopts a speculative research path combining normative analysis and policy text analysis. Normative analysis is based on the ethical theory of educational leadership, and makes value judgment and normative attribution of ethical risks. The policy text analysis systematically sorts out the relevant policy documents of the state and the Ministry of Education, the AI teaching management norms of universities and the reports of international organizations, and provides institutional context support for normative judgment. The overall analysis framework follows the three-stage logic of " risk identification-ethical attribution-institutional response. " Firstly, the typical ethical risk scenarios are summarized, and then the normative attribution is carried out with the three-dimensional ethical framework. Finally, the institutional response strategy is derived based on the attribution conclusion.

2. Literature review and theoretical basis

2.1 Research progress on the integration of generative artificial intelligence and university teaching

In recent years, with the rapid iteration of generative artificial intelligence tools such as ChatGPT and Wen Xin Yi Yan, the academic community has carried out multi-dimensional discussions on its integration with college teaching, and the relevant research can be roughly sorted out along three lines. The first is functional research, focusing on the technical support role of generative artificial intelligence in teaching design, resource development, formative evaluation and other links, pointing out that it can reduce the burden of teachers ' routine work and improve teaching efficiency. [3] The second is the effect study, which focuses on the impact of artificial intelligence-assisted teaching on students ' learning engagement, academic achievement and personalized learning experience. Some empirical studies reveal that its effect is subject to subject characteristics and usage methods. [4] The third is risk research, involving academic integrity, data privacy, algorithm bias and other issues, but most of them stay in the description of phenomena and the call for countermeasures. [5] In general, the existing research has accumulated a lot in the technical efficiency dimension, but the normative analysis of the ethical dimension is still weak, especially the lack of systematic review from the perspective of educational leadership ethics, which leaves room for the theoretical entry of this paper.

2.2 The theoretical pedigree of educational leadership ethics

Educational leadership ethics provides normative theoretical resources for examining technology-enabled teaching. Caring ethics is represented by Noddings, emphasizing that the core of educational relationship lies in 'caring ' rather than 'efficiency '.Teaching leadership behavior should take the establishment and maintenance of caring relationship as the ethical purpose. [6] Justice ethics is based on the principle of justice of Rawls, which requires the allocation of teaching resources and institutional arrangements to follow the

principle of fairness, with particular attention to the protection of the interests of vulnerable groups. [7] Responsibility ethics originates from Weber 's distinction between " responsibility ethics " and " belief ethics, " emphasizing that actors must be responsible for the foreseeable consequences of their decisions, which provides an ethical basis for multi-agent accountability. [8] The three together constitute the analytical pedigree of educational leadership ethics in this study.

3. Research Design

3.1 Orientation of research methodology: normative analysis and policy text analysis

This study is a theoretical study of speculative orientation, with normative analysis as the core method, supplemented by policy text analysis. This choice stems from the normative nature of the research question. The topics such as the elimination of academic integrity, the materialization of teacher-student relationship, and the solidification of algorithmic bias caused by the teaching of generative artificial intelligence empowering colleges and universities are not purely technical efficiency issues, but rather value judgments involving ' how teaching should be '. It is not possible to rely solely on empirical description or causal explanation, but to conduct normative argumentation under a clear ethical framework. Normative analysis provides a logical tool for examining practical practice and deducing ought-to-be institutional arrangements from ethical principles. Policy text analysis is used to examine the ethical presupposition and value orientation contained in the artificial intelligence education policies issued by the state and education authorities, and to provide institutional context and text basis for normative analysis. The combination of the two not only ensures the theoretical depth of ethical argumentation, but also makes the institutional response strategy realistic.

3.2 Data and text sources

The analysis data of this study mainly cover three categories : First, policy documents at the national and Ministry of Education levels, including the " New Generation of Artificial Intelligence Development Action Plan (2025-2029) " " Artificial Intelligence + Education " Action Plan, " etc., to grasp the basic position and normative orientation of the country 's application of artificial intelligence education ; secondly, the normative texts of artificial intelligence teaching management at the university level, including the guidelines for the use of AI tools issued by some universities, the provisions on generative artificial intelligence in academic integrity regulations, etc., are used to examine the current situation and gap of system supply. Thirdly, representative academic literature and international organization reports, such as UNESCO 's " Guidelines for Education and Generative Artificial Intelligence, " are used to obtain empirical references for international comparative perspectives and ethical governance. The three types of data together constitute the text analysis corpus of this study, which provides a textual analytical basis for ethical attribution and institutional derivation.

3.3 Analysis path: risk identification-ethical attribution-institutional response

This study follows the progressive logic chain of ' risk identification-ethical attribution-institutional response ' to demonstrate. Firstly, based on the policy text and the actual teaching scene, this paper systematically peels off the multiple ethical risks of generative AI embedded in the teaching process of colleges and universities, clarifies the external representation and practical harm of various risks, and completes the problem diagnosis at the phenomenon level. Secondly, relying on the ethical framework of educational leadership, this paper explores the deep motivation behind the risk from the three ethical dimensions of care, justice and responsibility, explains how ethical contradictions are generated and amplified in the application of technology, and realizes the transition from phenomenon description to theoretical attribution. Finally, combined with the previous risk analysis and ethical interpretation, based on the context of China 's education policy, this paper deduces the appropriate institutional response plan, and puts forward the ethical governance strategy that can be implemented. The three paths are interlocking, realizing the unity of fact description, ethical interpretation and normative construction, and ensuring the self-consistency of theoretical deduction logic.

4. Ethical risk identification of generative artificial intelligence empowering college teaching

The deep embedding of generative artificial intelligence in college teaching is not simply a technical upgrade, but an ethical practice involving knowledge production, relationship reconstruction and power redistribution. The efficiency dividend often conceals its value conflict and ethical tension. Without systematic identification, the " enabling " narrative can easily slide into the " negative " reality. This chapter identifies its core ethical risks from four dimensions : academic integrity, teacher-student relationship, algorithmic bias and teaching autonomy.

4.1 Academic integrity risk: the attribution ambiguity and evaluation distortion of generated content

The impact of generative artificial intelligence on academic integrity is first manifested in the blurred boundary of content attribution. When students use AI to complete homework, papers and even exams, the generated content is neither ' other people 's work ' nor independent original, but a mixed product of human-computer interaction, which directly shakes the legitimacy of academic evaluation based on the default premise of ' students ' independent completion '. If there is a lack of AI use declaration system and content traceability mechanism, it is difficult for teachers to judge students ' real learning effectiveness, and the fairness of evaluation will be impaired. If the content generated by AI is heavily infiltrated in dissertations without being identified, the credibility of degree awarding will face systematic erosion. Academic integrity is related to the basis of trust on which the academic community depends. Generative artificial intelligence creates cracks in this foundation.

4.2 the risk of teacher-student relationship: the substitution and dilution of human-computer interaction for teaching care

The essence of teaching is not only the transfer of knowledge, but also the spiritual encounter between teachers and students with care as the link. The wide application of the generative artificial intelligence tutoring system is partially replacing the teacher 's feedback, question answering and guidance functions, causing the risk of weakening the emotional connection between teachers and students. When students are accustomed to seeking instant answers and personalized feedback from AI, the frequency and depth of face-to-face communication between teachers and students are bound to decline, and ' caring encounter ' is facing the dilemma of being diluted by technical interaction. Although AI can provide accurate knowledge response, it cannot perceive students ' emotions, understand their learning anxiety or give sincere encouragement. In the long run, teaching may degenerate into a process of de-emotional information transmission, students are simplified as data nodes, and teachers are reduced to assistants of technical systems. The alienation of this relationship weakens the educational function of education and deviates from the essential provisions of teaching as ethical practice.

4.3 Algorithm bias risk: the erosion of training data bias on teaching fairness

The output of generative artificial intelligence is deeply embedded in the distribution structure of training data, carrying social prejudice, cultural prejudice and gender stereotypes. When prejudice is solidified by models and enters teaching, it constitutes a hidden erosion of fairness, which is manifested in three aspects : excessive presentation of a certain culture or group experience in content generation, marginalization of other knowledge traditions ; in the evaluation feedback, non-standard expressions or minority language habits are given a systematic low evaluation ; in the recommendation system, disadvantaged students are guided to low-level resources and education stratification is solidified. Algorithmic bias has the characteristics of concealment and black box. It is often embedded in teaching decision-making with the appearance of ' objective neutrality ', which makes it difficult to identify and account for injustice. Justice ethics requires institutional arrangements to maximize the benefits of the least beneficiaries, while algorithmic bias weakens the right of vulnerable groups to obtain fair educational opportunities under the cover of efficiency, which constitutes a structural threat to teaching justice.

5. Risk attribution from the perspective of educational leadership ethics

5.1 Ethical dimension of care: the masking effect of technical efficiency logic on relational care

Noddings ' caring ethics regards education as the establishment and maintenance of ' caring relationship ', and emphasizes that teachers ' acceptance, response and empathy are the ethical foundation of teaching activities. In this way, we examine the teaching embedding of generative AI, and its efficiency orientation is systematically masking this relational nature. AI 's involvement in lesson preparation, question answering and feedback improves the speed and accuracy of information transmission, but simplifies the teacher-student interaction with emotional communication as a technical process of " question-response. " When the AI tutoring system replaces the teacher 's immediate response, teaching is degraded from ' encounter ' to ' service delivery '. Students get answers rather than understanding, and teachers give instructions rather than care. This substitution is a structural dissolution of the relationship between teaching ethics-the technical system can only simulate the response and cannot generate a real caring relationship. The most ethical ' presence ' in teaching is digested by efficiency logic, and the trust bond between teachers and students is at risk of breaking.

5.2 Ethical dimension of justice: algorithm power asymmetry and imbalance of teaching resource allocation

Rawls ' principle of justice requires institutional arrangements to maximize the benefits of the least beneficiaries and to ensure fair and equal opportunities. Examining the reality of generative AI empowering college teaching can reveal the dilemma of triple justice. First, the technology access gap between institutions. The gap between key universities and local colleges in computing resources, technology platforms and data accumulation has significantly differentiated the quality of AI teaching, and it is difficult for students in disadvantaged colleges to obtain equal support. Second, the difference in algorithm literacy between teachers and students. Some teachers are able to master AI and critically examine its output, while others passively rely on system recommendations due to insufficient technical ability, resulting in homogenization of instructional design and weakening students ' opportunities for diversified education. Third, structural bias in training data. AI output implies gender, race and cultural prejudice, and is embedded in the teaching process through content recommendation and evaluation algorithms, which constitutes implicit injustice to minority students. The asymmetric distribution of algorithmic power makes technology empowerment possible to solidify or even aggravate the inequality of teaching resources and opportunities.

5.3 The ethical dimension of responsibility: the fracture and ambiguity of the multi-agent accountability chain

Weber 's ethics of responsibility emphasizes that actors are responsible for the foreseeable consequences of their decisions. Generative artificial intelligence empowers college teaching to involve multiple subjects such as developers, platforms, managers, teachers and students. The responsibility chain should be linked, but it is broken and blurred in reality. Developers evade responsibility with algorithm black box and technology neutrality ; the platform borrows the disclaimer to pass the consequences on to universities and teachers ; the supply of management system lags behind, and AI teaching norms and accountability mechanisms are absent ; teachers ' responsibility cognition is not clear, and they lack the ability to judge the compliance of AI output ; students ' sense of responsibility is weak, often directly submit the generated content and ignore academic integrity. Multi-agent participation does not bring reasonable sharing of responsibilities, but leads to ' organized irresponsibility '. This fracture stems from the black box of technology, the lag of system renewal and the dispersion of subject interests. When ethical problems occur, all parties prevaricate each other, and ultimately damage the teaching ethical order and students ' rights and interests. Responsibility ethics not only requires accountability after the event, but also emphasizes foresight and process responsibility in advance. It requires colleges and universities to clarify the boundary of rights and responsibilities of each subject in the whole chain of ' design-deployment-use-feedback ', and establish a traceable and attributable accountability mechanism, so that ethical responsibility has a clear undertaker in each link of the technology life cycle.

6. Institutional response: university teaching governance strategy based on educational leadership ethics

The identification and attribution of ethical risks must ultimately be implemented as operational institutional arrangements to avoid ethical analysis staying at the level of theoretical speculation and deviating from the actual operating scenarios of college teaching. Based on the normative attribution of the aforementioned three-dimensional ethical framework of educational leadership, this paper puts forward the institutional response strategy of university teaching governance from the four levels of ethical regulation, organizational system, teaching governance and capacity building, and tries to realize the paradigm shift from ' remedial ethics ' to ' preventive ethics ', promote the construction of a teaching operation system that is compatible with technological empowerment and ethical constraints, and hold the fundamental educational position of moral education in the intelligent era.

6.1 Ethical regulation level: constructing a preventive AI teaching ethics system

Preventive ethical regulation emphasizes the establishment of behavioral boundaries before the application of technology. First of all, colleges and universities should formulate special ethical guidelines for the use of AI teaching, and clarify the allowable scope and prohibitions of generative artificial intelligence in the preparation, teaching, homework, evaluation and other links. Secondly, it is necessary to establish an ethical review mechanism for AI teaching applications, conduct an ethical risk assessment of AI tools to be introduced into teaching, and focus on reviewing the legitimacy of their data sources, algorithm transparency and potential bias. Thirdly, it is necessary to clarify the red line of academic integrity, establish an AI use declaration system, require students to truthfully mark the scope and method of use when submitting AI-assisted assignments, and shift the maintenance of academic integrity from " post-punishment " to " process transparency. "

6.2 Organizational system level: establish a multi-agent collaborative teaching ethics accountability mechanism

The implementation of responsibility ethics depends on a clear division of rights and responsibilities and an effective accountability chain. Colleges and universities should set up a school-level AI teaching ethics committee to absorb the participation of technology developers, teaching managers, front-line teachers, student representatives and off-campus ethics experts to form a multi-negotiated governance structure. On this basis, it is necessary to clarify the boundary of rights and responsibilities of each subject : the developer is responsible for the fairness and transparency of the algorithm, the manager is responsible for the system supply and resource guarantee, the teacher is responsible for the ethical judgment in the teaching process, and the student is responsible for the integrity of their own learning behavior. Therefore, a ' design-deployment-use-feedback ' full-chain accountability system is constructed, so that ethical responsibility can be traced and attributable to each link of the technology life cycle.

6.3 Teaching governance level: Reconstruct the teaching decision-making model with human-machine collaborative leadership

The promotion of human-machine collaborative leadership depends on breaking the dual thinking of technology dominance or human arbitrariness and reshaping the decision-making logic of teaching activities. Colleges and universities should rationalize the functional orientation of human-machine in the teaching field, establish the basic principles of teachers ' professional judgment at the core and generative artificial intelligence as an auxiliary tool, and incorporate AI analysis results into teaching decision-making reference rather than directly as a decision-making basis. On this basis, the boundaries of the respective roles of human and machine are clarified : artificial intelligence undertakes transactional work such as data collation, data statistics, and preliminary drafting of programs, and teachers lead curriculum design, academic research and judgment, value guidance, and educational decisions. Therefore, the teaching decision-making process of " man-machine research and judgment-teacher decision-dynamic adjustment " is built to avoid the squeeze of algorithm on the autonomy of teaching specialty, so that the goal of technical service and education can be

achieved, and the organic coupling of teachers ' professional wisdom and the advantages of artificial intelligence tools can be realized.

6.4 The level of capacity building: to build a two-way AI teaching ethics literacy cultivation system between teachers and students

The effective implementation of ethical regulation, organizational accountability and human-machine collaborative governance relies on the comprehensive quality of AI ethics of teachers and students. Only by consolidating the foundation of the main body 's ability can the hidden ethical risks of intelligent teaching be eliminated from the root and the ' last mile ' problem of system implementation be solved. Colleges and universities need to abandon the single and indoctrinating training mode, carry out the special training of AI teaching ethics regularly, build a two-way cultivation system covering teachers ' teaching leadership ability and students ' self-discipline ability, and improve teachers ' on-the-spot ethical decision-making ability in teaching scenarios through case discussion, teaching salon, man-machine collaborative teaching case review and other forms. For students, we should carry out digital academic integrity guidance, and jointly realize the all-round improvement of ethical cognition, ethical judgment and ethical practice ability.

7. Conclusion

Based on the normative analysis framework of educational leadership ethics, this paper systematically examines the ethical risks caused by generative artificial intelligence empowering college teaching and its institutional response path. The research shows that the risks of academic integrity digestion, teacher-student relationship materialization, algorithm bias solidification and teaching autonomy erosion are not accidental by-products of technology application, but the concentrated manifestation of educational leadership ethics in the intelligent era. The concealment of caring ethics, the imbalance of justice ethics and the rupture of responsibility ethics together constitute the deep ethical mechanism of risk generation. Accordingly, the institutional response should be based on the ethics of educational leadership, and should be promoted at the four levels of ethical regulation, organizational system, teaching governance and capacity building. ' Preventive ethics ' should be used to replace ' remedial ethics ', and ' human-machine collaborative leadership ' should be used to reconstruct the pattern of teaching governance, so as to achieve the dynamic balance between technology empowerment and ethical protection. This study is mainly based on normative analysis and policy text analysis, and there is still a lack of systematic empirical test. The practical applicability of the conclusions needs to be further verified. Future research can focus on the implementation effect of AI teaching ethics norms in colleges and universities, the empirical investigation of teachers and students ' ethical perception, and the case comparison of human-computer collaborative teaching leadership, so as to promote the theoretical deepening and practical implementation of educational leadership ethics in the context of intelligent technology.

Reference

[1] 刘明,郭烁,吴忠明,等.生成式人工智能重塑高等教育形态:内容,案例与路径[J].电化教育研究, 2024, 45(6):57-65.

[2] 郑煌杰.生成式人工智能的伦理风险与可信治理路径研究[J].科技进步与对策, 2025(12).

[3] 陈小强,肖丽萍,张慧.生成式人工智能赋能高职个性化学习路径研究[J].计算机应用文摘, 2024, 40(17):57-60.

[4] Berghel H , Berleant D , Berghel H .GenerativeAIin Higher Education: The Good, the Bad, and the Ugly[J].Computer, 59[2026-10-10].DOI:10.1109/MC.2026.3705552.

[5] 王树西,夏增艳.算法偏见,隐私与自主性:人工智能伦理困境破解路径研究[J].集成技术, 2025(5).

[6] 马香莲,杨绘.课堂教学边缘人的消解——基于诺丁斯教育关怀理论的视角[J].信阳师范学院学报：哲学社会科学版, 2019, 39(4):5.DOI:10.3969/j.issn.1003-0964.2019.04.014.

[7] 罗云,钟景迅,曾荣光.进城务工人员随迁子女教育公平问题的分配正义与关系正义之考察[J].北京大学教育

评论, 2015, 13(2):22.DOI:10.3969/j.issn.1671-9468.2015.02.011.
[8] 张甲子.信念·责任·学术——基于马克斯·韦伯伦理观的学术透视[J].继续教育研究, 2019(4):6.DOI:10.3969/j.issn.1009-4156.2019.04.006.