Scientific Research Publishing

Open Journal of Social Sciences, 2026, 14(9), 728-738
https://www.scirp.org/journal/jss
ISSN Online: 2327-5960
ISSN Print: 2327-5952

From "Technological Empowerment" to "Technological Erosion": Strategies for Protecting Critical Thinking in LLM-Assisted Instruction

Na Li, Ying Tian

School of Foreign Language, Northeast Electric Power University, Jilin, China
Email: lena1981_2000@126.com

How to cite this paper: Li, N., & Tian, Y. (2026). From "Technological Empowerment" to "Technological Erosion": Strategies for Protecting Critical Thinking in LLM-Assisted Instruction. *Open Journal of Social Sciences, 14,* 728-738.
https://doi.org/10.4236/jss.2026.149042

Received: August 24, 2026
Accepted: September 25, 2026
Published: September 28, 2026

Abstract

While large language models (LLMs) are reshaping instructional forms and realizing "technological empowerment", they also harbor risks of cognitive outsourcing, algorithmic dependence, and process black-boxing, which constitute forms of "technological erosion". These risks directly weaken learners' cognitive engagement and undermine the critical thinking that takes reasoning, questioning, and reflection as its core. In view of this, this paper analyzes the internal mechanism and main logic of the weakening of critical thinking in LLM-assisted instruction from the critical perspective. It further explores and constructs strategies for protecting critical thinking that can move from passive prevention to active immunity and from singular "empowerment" to a balance of "empowerment and protection".

Keywords

Large Language Models, Critical Thinking, Protection Strategies

1. Introduction

The explosive expansion and evolution of large language models (LLMs) are triggering a systemic reshaping of the higher education ecology. Generative artificial intelligence, represented by ChatGPT, DeepSeek, Gemini and other AI engines, has penetrated the entire educational chain—including instructional design, resource development, tutoring intervention, and assessment feedback—by virtue of its remarkable capacities in cross-modal generation, complex logical evolution, and large-scale knowledge reconstruction.

Many studies point out that the current digital transformation is confronted with serious challenges such as the erosion of technological ethics, blurred application boundaries, and uncertain prospects for human-AI collaborative models. University faculty and students are expected not only to attend to disciplinary knowledge but also to cultivate higher-order thinking abilities, among which critical thinking, creativity, communication, and collaboration are central. As a core quality of higher education, critical thinking is essentially "a process of inquiry and empirical verification grounded in an open rational spirit" (Dong, 2026), emphasizing a complete cognitive loop of questioning, analysis, argumentation, and reflection. However, the fluent, immediate, and highly structured information output provided by generative AI has led the convenience of technological tools to cross the boundary of assistance and replace human evaluation, logical inference, and viewpoint generation, thereby substantially compressing or even displacing the cognitive process of inquiry. Many empirical studies indicate that over-reliance on generative AI may induce intellectual inertia in learners (Qi & Zheng, 2026). When standard answers and formulaic knowledge can be instantly obtained through AI as an "external knowledge extension", the traditional value of knowledge assessment is rapidly deconstructed.

Technological empowerment represents an extension of the breadth of human cognitive tools, whereas technological erosion, including the narrowing of communication depth by information responses (Kou & Liu, 2025), the weakening relationship of socialization (Li & Song, 2026) points directly to the implicit decline and structural contraction of thinking subjectivity. Research on protecting critical thinking in LLM-assisted instruction is therefore both a theoretical return to the essence of education in the intelligent era and a realistic response to urgent pedagogical dilemmas.

2. Literature Review

This section presents a narrative synthesis of existing literature concerning the pedagogical impact of generative AI. While early studies focused primarily on efficiency and learning outcomes, recent scholarship increasingly highlights the unintended consequences of AI integration on student cognitive engagement. Drawing from these investigations, the primary crises associated with AI-assisted instruction—characterized as dimensions of "technological erosion"—can be synthesized as follows.

2.1. The Superficialization of Thinking: From "Efficiency Dependence" to "Cognitive Offloading"

Generative AI provides learners with unprecedented rapid access to information, often shifting the pedagogical focus toward mere efficiency and task completion. A substantial body of research demonstrates that the unmediated use of LLMs significantly diminishes learners' active engagement in information evaluation, reasoning, and reflection. Specifically, as learners outsource information retrieval,

preliminary judgment, and argument drafting to AI tools, process-oriented cognitive skills—such as information verification, argument construction, and source evaluation—experience severe attenuation (Zhang, 2026). Over time, this impedes the formation and transfer of critical thinking capacities (Li & Wang, 2026). This phenomenon, conceptualized as "cognitive offloading", poses a systemic risk of competence superficialization. Without metacognitive guidance and rigorous task constraints, learners frequently degenerate into "passive receivers" of immediate, well-organized answers, compressing multi-round inquiry processes into superficial "prompt games". Repeatedly delegating complex syntactic construction and logical inference to LLMs plunges deep cognitive engagement into a state of chronic dissipation, eroding the rigor of thought. Crucially, the amalgamation of hallucination generation, algorithmic bias, and data contamination within LLMs (Dai & Cai, 2026; Duan & Zhang, 2026; Sun et al., 2026) can induce critical failure and the emergence of "pseudo-cognition" among learners who lack adequate discriminative capacity.

2.2. The Risk of Information Credibility: From "Blind Trust" to the "Perils of Hallucination"

Faced with the severe cognitive risks posed by LLMs—particularly the crisis of information credibility where learners blindly trust AI outputs—early research largely adopted a "passive defense" approach. To mitigate the perils of AI hallucinations, institutions initially relied on prohibitive policies, rule-based detection, and academic integrity constraints (Guo & Huang, 2026; Chen & Chen, 2025). However, recognizing that passive defense fails to address the root cause of epistemic blind trust, a growing body of research has shifted toward a "collaborative empowerment" strategy. This paradigm acknowledges the profound risk of information bias and aims to reconstruct the learner's epistemic vigilance. Scholars advocate positioning AI as a guided "co-worker", emphasizing task restructuring, evidence transparency, and metacognitive training to enhance learners' independent judgment and source verification capabilities (Gu et al., 2026; Ma & Cui, 2026). Moving beyond mere rule-based control, this literature underscores the necessity of treating AI outputs as reflective triggers—compelling learners to verify sources, cross-examine arguments, and construct robust evidence chains (Hu et al., 2026; Sang et al., 2026). The essential principle is to counteract blind trust by repositioning AI from an omniscient "answer provider" to a "Socratic sparring partner", thereby mitigating hallucination harms and actively cultivating critical thinking.

2.3. The Dilemma of Monologic Interaction: From "Tool Avoidance" to "Passive Dependence"

The explosive integration of generative AI has inadvertently trapped pedagogical practices in a dilemma of monologic interaction. Initially, driven by concerns over hallucination risks and ideological biases, institutions largely resorted to "tool avoidance" or strict restriction strategies (Dai & Cai, 2026). Yet, practice demon-

strates that such avoidance is unsustainable and stifles personalized instructional innovation (Weng, 2025; Yuan, 2024). Consequently, as AI usage became inevitable, learners frequently developed a "passive dependence" on the technology's one-way question-and-answer model, severely limiting the breadth and depth of dialogic thinking. To break this monologic dependency, recent literature advocates defining AI as a controlled tool within adaptive governance frameworks, utilizing transparency mechanisms to exploit AI's capacity for differentiated feedback (Li & Song, 2026). Although personalized generation has shown promise in higher education, it remains hampered by technical biases and passive user engagement (Duan & Zhang, 2026). Thus, scholars increasingly champion "metacognitive co-training"—a triadic collaboration among teachers, students, and AI—as the ultimate antidote to passive dependency. Studies indicate that embedding explainable evidence-chained interfaces and utilizing AI for real-time, counter-perspective feedback can transform monologic acceptance into multidimensional cognitive conflict (Hu et al., 2026; Yi & Xie, 2023; Sang et al., 2026). Grounded in teacher-led curriculum redesign (Guo & Chen, 2023; Li & Song, 2026), this co-training paradigm requires the systematic integration of interaction logs and metacognitive scaffolding (Guo & Huang, 2026; Ma et al., 2026), marking a vital shift from one-way tool dependence to the dynamic, interactive construction of cognition.

2.4. The Dissolution of Learner Agency: From "Human-Machine Substitution" to "Value Erosion"

The reshaping of the educational landscape by generative AI compels the academic community to critically re-examine the boundaries of learning agency. A prominent crisis highlighted in current literature is the dissolution of learner subjectivity, largely driven by the "human-machine substitution" paradigm. While this substitution model yields short-term efficiency gains, its long-term mechanisms precipitate the alienation of learner agency, the blunting of creativity, and the complete instrumentalization of educational goals. Unchecked algorithm-driven ideological generation inherently risks profound value erosion and ideological infiltration (Wu & Gong, 2026; Yao & Cui, 2023). To counter this existential threat to educational ethics, the academic community widely proposes "human-machine symbiosis" as an ethical and sustainable alternative. The central tenet is to deconstruct AI's role as a cognitive "surrogate" and reconfigure it as a bounded "collaborator" (Gu et al., 2026; Ma & Cui, 2026). Implementing this symbiosis requires multi-layered scaffolding. At the micro-design level, intelligent agents must be engineered to reveal evidential uncertainty and catalyze deliberation (Hu et al., 2026; Yi & Xie, 2023). At the macro-organizational level, institutional data-governance frameworks, ethical review systems, and academic assessment standards must be rigorously reconstructed (Zhao & Dong, 2026). Empirical evidence confirms that under ethically demarcated human-machine divisions of labor, AI can absorb lower-order retrieval tasks, emancipating learners to immerse themselves

in higher-order meaning-making. Ultimately, this safeguards human value alignment and fortifies learner autonomy against the encroaching risks of technological erosion (Duan & Zhang, 2026; Ma et al., 2026).

3. Strategies for Protecting Critical Thinking

Based on the above problems and problem-related studies in Section 2, this section proposes four mutually reinforcing strategies to achieve the sustained protection and cultivation of critical thinking in instructional settings.

3.1. Evaluation Reconstruction Based on a "Process-Oriented View of Knowledge": From "Result Orientation" to "Logical Progression"

LLMs' instantaneous outputs risk reifying knowledge into disembodied products, directly triggering the "efficiency dependence" and cognitive offloading discussed earlier (Dong, 2026). To counteract this superficialization, instructional evaluation requires a fundamental epistemological shift: from an "entitative view" focusing on final answers to a "process-oriented view" emphasizing knowledge generation, verification, and reflection (Zhang, 2026). This reconstruction safeguards critical thinking by rendering AI-assisted cognitive processes both "visible" and "measurable".

To promote productive cognitive engagement aligned with learning objectives, educators must redesign evaluation rubrics through three core strategies: (1) Assessing "thinking traces". Formative evaluations should focus on the quality and evolution of students' AI-assisted reasoning processes, as reflected in prompt design and iterative refinement. When feasible, students may submit complete AI interaction logs documenting their prompt evolution. For students who cannot submit complete logs due to privacy or technical constraints, instructors may accept privacy-preserving alternatives, such as desensitized interaction summaries, structured reflective memos describing key decision points and revisions, or brief oral accounts of the reasoning trajectory. Regardless of format, evaluation criteria shift from mere answer accuracy to the logic, coherence, and depth of inquiry demonstrated in the documented process. (2) "Reverse-engineering" AI logic: Rather than passively consuming LLM-generated texts, students must actively deconstruct them. By mapping output structures and identifying conceptual leaps or reasoning gaps, this approach rigorously evaluates their logical identification capacities. (3) Evaluating "cognitive revision": Establishing "error rectification" as an independent scoring dimension encourages students to intentionally prompt LLMs for flawed solutions and subsequently refute them.

Through this process-oriented paradigm, LLMs transition from cognitive "subcontractors" into instructional "scaffolds". Ultimately, this mechanism forcibly disrupts the pathway to cognitive offloading, sustaining the high-intensity intellectual engagement requisite for critical thinking.

3.2. Instructional Design Embedded with "Hallucination Identification": From "Passive Reception" to "Evidence Verification"

To address the epistemic crisis where learners blindly trust algorithmic outputs (Dai & Cai, 2026), instructional design must evolve. Transforming the technical deficiency of AI "hallucinations" into a pedagogical opportunity requires shifting from mere content transmission to rigorous evidence-checking, anchored in "epistemic vigilance".

To institutionalize the analysis of AI fabrication and algorithmic bias, instructors should implement three cohesive operational strategies: (1) Deploying "fallacy anchors": Educators intentionally prompt LLMs to generate outputs infused with factual inaccuracies or logical fallacies. These synthesized errors serve as pedagogical "anchors", providing concrete negative examples that compel students to engage in rigorous, diagnostic fact-checking. (2) Applying discipline-specific evidence standards: Moving beyond arbitrary, one-size-fits-all quantitative mandates (such as a universal requirement for "at least two sources"), instructional design must be anchored in the epistemological norms of each discipline. Students must corroborate AI outputs using field-appropriate authoritative benchmarks. Furthermore, to explicitly assess the reliability of AI-generated citations, students must execute a rigorous tripartite verification protocol: (a) *authentication*, confirming the actual digital or physical existence of the reference to eliminate AI-generated citations; (b) *credibility evaluation*, scrutinizing the peer-review status and academic standing of the publication venue; and (c) *contextual cross-referencing*, ensuring the original source genuinely substantiates the specific claims attributed to it by the AI. (3) Documenting "algorithmic bias": Students should maintain metacognitive reflection logs to systematically track and analyze the AI's cultural, gender, or theoretical biases, thereby cultivating the critical capacity to deconstruct embedded value orientations.

Ultimately, institutionalizing this "hallucination identification" framework actively mitigates the risk of AI-induced "knowledge poisoning". It fundamentally awakens students' intellectual agency, forging resilient epistemic vigilance and independent judgment through a continuous doubt-verification cycle.

3.3. Collaborative Debate Mechanism from the Perspective of Intersubjectivity: From "One-Way Question-and-Answer" to "Multidimensional Conflict"

Traditional AI-assisted instruction is mostly confined to a dualistic subject-object model of one-way question-and-answer, in which students tend to regard AI as an omniscient authority. This singular interaction severely limits the breadth and depth of thinking. Introducing the perspective of intersubjectivity means treating the large language model as a "dialogue partner" with multiple viewpoints, thereby creating a complex triadic dialogue space of "human-machine-human" and using the multidimensional output of technology to stimulate deep cognitive conflict.

To reshape this interactive model, teachers can construct debate arenas with high tension: (1) Preset prompt-based contests among "multiple positions". When discussing controversial social issues or complex scientific ethics, instruct students to require the large model to play sharply contrasting roles (e.g., a "radical technological determinist" versus a "conservative ecological protectionist"), and ask students to locate logical ruptures between two diametrically opposed AI arguments, thereby training their ability to extract insight from contradiction. (2) Build a triadic debate space of "student-AI-peer". Design "human-machine paired" debate competitions, in which "Student A + AI assistant" confronts "Student B + AI assistant". Under this mechanism, students must not only refute the opposing human debater's claims but also identify and attack evidential weaknesses in the opposing AI assistant's arguments, thereby deepening critical thinking through multidimensional intellectual collision. (3) Implement "role reversal" Socratic questioning. Break the convention that students are always questioners. Use prompts to endow AI with the role of "critical examiner" (e.g., "Please offer the most rigorous rebuttal to my following argument"), forcing students to continuously refine their defensive logic and theoretical depth when confronted with AI's step-by-step pressure and logical questioning.

Under this collaborative debate mechanism, artificial intelligence is no longer the terminal that provides standard answers; rather, it becomes a catalyst that continuously generates cognitive conflict and extends the boundaries of thinking. Multidimensional collisions of perspective effectively break the limits of one-dimensional thinking and truly realize the leap from "tool substitution" to "human-machine symbiosis and collaborative empowerment".

3.4. Ethic Framework for Intelligent Instruction Based on "Value Alignment": From "Technology-Driven" to "Human-Centered Guidance"

The deep erosion caused by LLMs in education is often accompanied by technology's encroachment on educational ethics and on the value attributes of the human person. To fundamentally avoid "technological erosion", it is necessary to establish ethical regulations based on "value alignment" at both the macro and micro levels. Such regulations require that all applications of intelligent technology in teaching unconditionally submit to the core educational values of promoting human cognitive growth and safeguarding academic integrity.

Constructing this ethics framework requires translating abstract value concepts into executable operational norms: (1) Establish a tool-use contract based on the principle of "human-in-the-loop". At the initial stage of a course, teachers and students jointly sign an "AI-assisted learning ethics statement", clearly delimiting which cognitive processes (e.g., initial concept exploration and broad literature scanning) may draw on AI and which core processes (e.g., value judgment, core thesis generation, and expression of emotional resonance) must be absolutely dominated by the human mind, thereby drawing insurmountable "red lines" for tech-

nology use. (2) Implement output-review mechanisms based on "transparency and explainability". Establish a class-level "AI-generated content ethics review committee". When students submit works containing AI-assisted generation, they must actively declare the proportion of AI involvement and its main uses. AI outputs involving sensitive topics must undergo dual scrutiny through peer evaluation and teacher review to ensure that the content does not violate mainstream values or academic bottom lines. (3) Cultivate a digital literacy ecology of "technological responsibility awareness". Make "AI ethics and intellectual property" an important component of general education, guiding students to explore deeper issues such as data privacy, knowledge exploitation, and algorithmic hegemony, so that they are transformed from "blind consumers of technology" into "rational users with digital responsibility".

The ethics framework based on value alignment constitutes the last line of defense against technological erosion. At both institutional and cultural levels, it safeguards the "good-oriented" application of technology and ensures that, in the era of human-machine symbiosis, the control of education remains firmly in the hands of human beings with independent thinking and value judgment.

4. Conclusion

Large language models first entered the instructional arena, carrying the optimistic expectation of "technological empowerment". However, when generative efficiency replaces cognitive processes, answer supply obscures evidential responsibility, and fluent expression dissolves deep thinking, technological application may slide from "empowerment" to "erosion". The critical reflection of this article shows that the risk does not stem from technology itself but from the passive adaptation of the instructional system to intelligent tools under the logic of efficiency. When assessment attends only to outcomes, when curricula ignore verification, when interaction stops at question-and-answer, and when ethics yields to instrumental rationality, students' critical thinking quietly degenerates in the midst of convenience. Therefore, the key to pedagogical response lies not in simply rejecting or restricting AI, but in rebuilding an instructional structure and ethical order led by human values so that technology truly serves the development of higher-order thinking.

On this basis, the present article proposes four instructional strategies centered on protecting critical thinking: reconstructing assessment through a process-oriented view of knowledge so as to move learning from "result orientation" back to "logical progression"; embedding hallucination identification and evidence verification into curricula so that learners shift from "passive reception" to "active verification" of AI output; constructing a human-machine collaborative debate mechanism from the perspective of intersubjectivity so that one-way question-and-answer is transformed into multidimensional collision; and building an ethics framework for intelligent instruction based on value alignment so that all stages from prompt construction to output review remain under the scrutiny of educa-

tional value. These four strategies respond respectively to systemic imbalances in assessment systems, curriculum content, interaction structures, and ethical norms, and jointly point to a single goal: to transform large language models from "answer terminals" that replace cognition into "thinking intermediaries" that stimulate cognitive conflict, strengthen evidential responsibility, and expand dialogic space.

It should be noted that the discussion in this article is primarily a theoretical-critical analysis and strategic construction, and it has not yet systematically examined the actual effects of the above strategies across different disciplines, educational stages, or technological contexts through empirical research. At the same time, large language models are iterating rapidly, and the continuously changing boundaries of their capabilities will keep rewriting the forms of teaching risks and responses. Future research should conduct design-based studies and longitudinal assessments in real instructional settings to further investigate the operability of these strategies and their long-term impact on the development of students' critical thinking. In addition, the concrete realization of value alignment in education involves multiple issues such as technology ethics, curriculum governance, and teacher professional development, which also require sustained interdisciplinary inquiry.

Ultimately, whether large language models can truly serve education depends on whether educators can maintain a clear and sustained tension between the logic of technology and the logic of human cultivation. Technological empowerment should not become a new variant of efficiency worship, nor is technological erosion an irreversible fate. Only by placing the protection of critical thinking at the center of intelligent instructional design and by insisting that every embedding of technology be guided by human values can artificial intelligence be elevated from a convenience tool to a dialogic force that promotes deep learning. This is the academic concern of the present study, and it is also the fundamental proposition to which teaching reform must respond in the age of intelligence.

Funded Project

Higher education research project of Jilin Institute of Higher Education (JGJX2023C31).

Author Contributions

Li Na and Tian Ying jointly contributed to Writing—original draft and Writing—review & editing. Li Na additionally conducted the Formal analysis and was responsible for Project administration and Funding acquisition. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

Chen, C. L., & Chen, L. Y. (2025). Theoretical Basis, Risks, and Response Strategies for Integrating Artificial Intelligence into Ideological and Political Education in Higher Education "Dual Classrooms"—A Review Based on Marxist Theory. *Journal of Yunmeng, 46,* 92-99. (In Chinese) https://doi.org/10.16740/j.cnki.cn43-1240/c.2025.03.004

Dai, J. C., & Cai, Z. (2026). Generative AI Hallucination and the Post-Truth Risk. *Studies in Philosophy of Science and Technology, 43,* 101-107. (In Chinese) https://doi.org/10.3969/j.issn.1674-7062.2026.03.015

Dong, J. Q. (2026). From Cognitive Offloading to Cognitive Outsourcing: Rethinking Technology Empowerment and De-Intellectualization Issues. *Foreign Language and Literature, 42,* 159-172. (In Chinese) https://doi.org/10.26922/j.cnki.waiguoyuwen.2026.04.013

Duan, D. P., & Zhang, J. Q. (2026). Practical Pathways Challenges and Insights of AI-Supported Personalized Teaching in American Higher Education. *Heilongjiang Higher Education Research, 44,* 91-97. (In Chinese) https://doi.org/10.19903/j.cnki.CN23-1074/G.2026.08.013

Gu, X. Q., Cao, Y., & Ji, Z. Q. (2026). Design Principles, Technical Pathways and Learning Ecosystem Construction of Educational Agents from the Perspective of Human-AI Symbiosis. *Open Education Research, 32,* 28-37. (In Chinese) https://doi.org/10.13966/j.cnki.kfjyyj.2026.04.003

Guo, J. N., & Chen, W. Y. (2023). Security Risks of Generative Artificial Intelligence Technology and Its Prevention Strategies. *Think Tank of Science and Technology, No. 11,* 42-52. (In Chinese) https://doi.org/10.19881/j.cnki.1006-3676.2023.11.06

Guo, X. L., & Huang, W. (2026). Core Action Strategies for Faculty to Guide Students in the Standardized Use of Generative AI—A Study of Syllabi Policies on AI at Leading International Universities. *Studies in Foreign Education, 53,* 116-128. (In Chinese) https://doi.org/10.20250/j.sfe.2026.01.006

Hu, J. K., Tang, C., & Sang, G. Y. (2026). Generative Artificial Intelligence Empowers Digital Textbook Design: Value Clarification, Risk Review and Path Exploration. *Modern Educational Technology, 36,* 46-55. (In Chinese) https://doi.org/10.3969/j.issn.1009-8097.2026.07.005

Kou, Y. D., & Liu, X. (2025). Artificial Intelligence as the Other: Deconstruction and Reconstruction of the Structure of Teaching Interaction. *Contemporary Education Sciences, No. 6,* 12-22. (In Chinese) https://doi.org/10.3969/j.issn.1672-2221.2025.06.002

Li, J. J., & Wang, T. J. (2026). Four-Dimensional Orientations and Their Transcendence in Education's Response to Generative Artificial Intelligence. *Journal of Hebei University (Philosophy and Social Science), 51,* 108-119. (In Chinese) https://doi.org/10.3969/j.issn.1005-6378.2026.03.011

Li, X. J., & Song, Z. K. (2026). Actor Roles and Risk Governance in AI Agent-Driven Teaching Reform. *China Educational Technology, No. 8,* 79-86. (In Chinese) https://doi.org/10.3969/j.issn.1006-9860.2026.08.011

Ma, X. R., Li, M., & Chen, W. Y. (2026). Balancing Teacher Development in the Era of Generative Artificial Intelligence. *Journal of Architectural Education in Institutions of Higher Learning, 35,* 9-16. (In Chinese) https://doi.org/10.11835/j.issn.1005-2909.2026.03.002

Ma, Z. Q., & Cui, X. (2026). Human-Machine Collaborative Teaching Models from the Perspective of Intersubjectivity: Theoretical Framework and Practical Forms. *e-Educa-*

tion Research, 47, 14-21. (In Chinese) https://doi.org/10.13811/j.cnki.eer.2026.03.002

Qi, S. P., & Zheng, J. Y. (2026). Cognitive Offloading Risks of Generative Artificial Intelligence and Their Mitigation. *Journal of Jishou University (Social Sciences Edition), 47,* 52-60+82. https://doi.org/10.13438/j.cnki.jdxb.2026.04.006

Sang, G. Y., Hu, J. K., & Wen, L. M. (2026). Generative Digital Textbooks Empowering New Instructional Paradigms: Value Representation, Potential Risks and Implementation Pathways. *e-Education Research, 47,* 121-128. (In Chinese) https://doi.org/10.13811/j.cnki.eer.2026.02.015

Sun, Y. Y., Wen, S. F., Su, J. Y., & Peng, D. (2026). Addressing the Risk of Cognitive Outsourcing: A Learning-by-Teaching Human-AI Collaboration Model Based on Scientific Argumentation Tasks. *Modern Distance Education Research, 38,* 103-112. (In Chinese) https://doi.org/10.3969/j.issn.1009-5195.2026.03.012

Weng, M. (2025). The Coupling Mechanism, Risk Manifestation and Countermeasures of DeepSeek's Involvement in Ideological and Political Education. *Hongyan Chunqiu, No. 10,* 114-120. (In Chinese) https://doi.org/10.16684/j.cnki.hycq.2025.10.024

Wu, Z. H., & Gong, Y. L. (2026). The Transformation of Higher Education in the Era of Generative Artificial Intelligence: Ontological Challenges, Alienation Risks, and Systemic Reconstruction. *Journal of Xiangtan University (Philosophy and Social Sciences), 50,* 110-116. (In Chinese) https://doi.org/10.13715/j.cnki.jxupss.20260519.001

Yao, J. J., & Cui, Y. H. (2023). From Deconstruction to Construction: The Generation Mechanism and Response Strategies of ChatGPT's Ideology. *Journal of the Party School of Tianjin Committee of the CPC, 25,* 53-62. (In Chinese) https://doi.org/10.16029/j.cnki.1008-410X.2023.06.006

Yi, X. Y., & Xie, X. (2023). Unpacking the Ethical Value Alignment in Big Models. *Journal of Computer Research and Development, 60,* 1926-1945. (In Chinese) https://doi.org/10.7544/issn1000-1239.202330553

Yuan, Z. (2024). The Systematic Risk of Generative AI on Education and the Regulation. *Journal of Political Science and Law, No. 6,* 46-59. (In Chinese) https://doi.org/10.3969/j.issn.1002-6274.2024.06.004

Zhang, L. (2026). Process-Oriented View of Knowledge: The Core Imperatives for Teaching Epistemology in the Age of Artificial Intelligence. *Journal of Educational Studies, 22,* 73-86. (In Chinese) https://doi.org/10.14082/j.cnki.1673-1298.2026.03.007

Zhao, G. H., & Dong, L. C. (2026). Factors Influencing Pre-Service Teachers' Acceptance of Artificial Intelligence Technology: An Empirical Analysis Based on the UTAUT Model. *Journal of Guangxi Normal University (Philosophy and Social Sciences Edition), 62,* 92-103. (In Chinese) https://doi.org/10.16088/j.issn.1001-6597.2026.03.010