Available Online: 13 May 2026 DOI: 10.54254/3049-7248/2026.33582

Shocks and adaptation: an international comparative study on the reconstruction of graduate supervision relationships in the AI era

Yuan Liu*, Meihui Chang

School of Educational Science, Xinyang Normal University, Xinyang, China

*Corresponding Author. Email: liuyuan1395@163.com

Abstract. The rapid advancement of Artificial Intelligence (AI) technology has evolved it from an efficiency tool into a technology of power that reconstructs power relations in education. Grounded in the Technology Acceptance Model (TAM) and Foucault's Power-Knowledge Theory, this study constructs a three-dimensional analytical framework of operation-supervision-institution. By comparing typical cases in China, the United States, and Russia, it systematically examines the collaborative system between AI and human supervisors. The findings reveal: (1) Efficiency improvements at the operational level are accompanied by the transfer of cognitive power between supervisors and students; (2) The supervisory level gives rise to a power struggle between algorithmic authority and the authority of human supervisors; (3) The institutional level reflects the cultural preservation efforts of various countries through normative institutionalization. Targeting China's "Supervisor as Primary Responsible Party" system, this study constructs a hierarchical and actionable three-tier collaborative framework, providing cross-national evidence for policy formulation.

Keywords: AI tutor, graduate supervision, power reconstruction, functional complementarity, international comparison, algorithmic authority

1. Introduction

The rapid development of artificial intelligence technology is driving the digital transformation of education to an in-depth level. Against this backdrop, the *Opinions on Accelerating the Digital Transformation of Education* (Jiao Ban [2025] No. 3), jointly issued by nine ministries including the Ministry of Education, explicitly proposes to "integrate artificial intelligence technology into all elements and processes of education and teaching" and "explore new models of human-AI collaborative teaching", marking that human-AI collaboration has been elevated to a key agenda in educational practice. This policy orientation not only aims to realize structural changes in the education system through technological empowerment to support the construction of a strong education nation, but also puts forward a profound demand for the reconstruction of inter-subject relations and power structures in traditional educational contexts, indicating that the education system will face a profound "shock" and urgently needs systematic "adaptation". In this process, the traditional role of supervisors is the first to be affected. Artificial intelligence has greatly expanded access to information and significantly enhanced students' autonomy and agency, making supervisors no longer the sole source of knowledge; the interaction model between supervisors and students has evolved from one-way knowledge transmission to a complex three-way collaboration among graduate students-supervisors-artificial intelligence. While this new pattern improves educational efficiency and openness, it also triggers severe shocks such as the disordered flow of power, the blurring of responsible subjects, and the deconstruction of traditional supervision relationships. Thus, a core research issue emerges: in the irreversible context of technological penetration, how should the education system "adapt" to reconstruct an effective balance of power, a clear accountability framework, and a sustainable collaborative system in graduate supervision?

Therefore, exploring how the education system implements effective adaptation to achieve a new balance has become an issue of both academic urgency and practical value. However, existing research presents a clear polarization: one side focuses on the micro-level utility of technology adoption but neglects to examine the changes in power relations caused by technology; the other emphasizes macro-level power criticism, yet its insights are difficult to translate into context-specific adaptation schemes. This theoretical divide results in a lack of a coherent and operable meso-level explanation for the "shock-adaptation" process. To bridge this analytical gap, this study adopts an analytical perspective centered on "shocks and adaptation" and constructs a three-dimensional analytical framework of operation-supervision-institution that connects micro-level behaviors and macro-level structures. Based on this framework, this paper seeks to answer the following questions: What forms do the "shocks" of artificial intelligence to graduate supervision relationships take at the operational, supervisory, and institutional levels respectively? How do China, the United States, and Russia—three countries with distinct governance logics—implement differentiated "adaptation" across these three dimensions, thereby providing theoretical reference and policy implications for China's graduate education, which is exploring its own path—namely, how to transform technological shocks into opportunities for educational innovation through proactive and systematic institutional design and cultural guidance.

2. Analytical tool: construction of the "functional stratification-power balances" three-dimensional model

In the field of higher education, especially graduate education centered on knowledge production, the strong intervention of AI tools (e.g., ChatGPT, Copilot) in literature retrieval, data analysis, and thesis writing has greatly improved the efficiency of academic research, yet it also poses profound challenges to supervisors' authority, academic ethics, and even the essence of education. However, current academic discussions on this phenomenon exhibit obvious theoretical fragmentation and dialog barriers, which are mainly reflected in two parallel research strands.

On the one hand, mainstream research focuses on the "usefulness" and "ease of use" of AI tools, with a core orientation of technological functionalism-driven efficacy verification. A large number of empirical studies aim to demonstrate the effectiveness of AI in specific educational scenarios, explore how it improves students' learning outcomes and provides personalized feedback, or verify the applicability of the Technology Acceptance Model (TAM) in explaining the willingness of supervisors and students to use AI [1-4]. While such research has important practical value and forms an empirical basis for technology adoption, its research agenda typically revolves around "how to integrate technology more effectively", often regarding social dimensions such as ethics and power as external constraints or secondary considerations for technology application rather than core research topics. This paradigm may lead research to "remain confined to superficial narratives of technological empowerment and efficiency enhancement" [5], ignoring profound changes in social relations and power structures behind technology application. On the other hand, critical research draws on power theories to raise structural doubts about technology at the macro level. Relevant studies criticize the erosion of teachers' professional autonomy by "algorithmic authority", reveal the tendency of "data colonialism" in data collection and use [6], or analyze how national AI strategies shape specific imaginaries of educational futures [7]. These critiques are highly enlightening, as they deconstruct macro power structures and propose ideal directions for reform (e.g., advocating democratic supervision and community participation). However, they lack mechanistic empirical analysis of how power transfer, struggle, and checks and balances occur in micro interactions (e.g., a specific thesis supervision session). As a result, while profound, such critiques may become abstract when divorced from specific practical scenarios and difficult to translate into actionable improvement schemes.

To bridge the theoretical gap between "micro-level technology adoption" and "macro-level power criticism", it is particularly necessary to construct an integrated analytical framework with both theoretical penetration and empirical explanatory power. The goal of the "functional stratification-power checks and balances" framework is precisely to bridge this divide and address a fundamental question: when supervisors and students readily adopt and use AI for its practical value, how are power relations in the field of graduate education subtly and profoundly reshaped? This research-driven shift in perspective focuses not only on "what utility AI provides" but also on "what relational changes the use of AI entails".

2.1. Theoretical origins: a combined perspective based on the technology acceptance model and power-knowledge theory

The theoretical foundation of the "functional stratification-power checks and balances" framework stems from the creative juxtaposition and integration of the Technology Acceptance Model (TAM) and Michel Foucault's Power-Knowledge Theory. In this framework, the two theories serve as the "initiation mechanism" and "deepening mechanism" respectively, collectively explaining the transformation of power structures after AI intervenes in educational relations.

As the initiation system in this framework, the core variables of the Technology Acceptance Model—perceived usefulness and perceived ease of use—provide a micro-foundation for understanding the initial dynamics of power flow [8]. The model indicates that the initial changes in power structures in the human-AI collaborative system stem not from external coercion but from rational choices made by actors—graduate students and supervisors—based on cost-benefit calculations. Students adopt AI because it improves efficiency in literature reviews, data processing, and code writing [9]; supervisors may also use AI to assist in handling repetitive supervision tasks. This voluntary adoption driven by perceived usefulness and ease of use constitutes the behavioral starting point of power transfer. Incorporating the Technology Acceptance Model into the analytical perspective grounds the research in concrete empirical behaviors, avoiding the grand narrative of technological determinism that regards technology as a unilaterally determining force.

If the Technology Acceptance Model explains the initial dynamics of power flow, Foucault's Power-Knowledge Theory, as the deepening mechanism, deconstructs individual technological choices at the level of social power structures, revealing profound changes behind voluntary power transfer. For Foucault, power is not only repressive but also productive, a force diffused in relational networks [10]. Accordingly, AI in education is not merely a neutral tool for improving efficiency but a contemporary technology of power that actively participates in knowledge production and power operation: it does not merely transmit knowledge but redefines it. It systematically reshapes knowledge production models and the subjectivity of supervisors and students through disciplinary mechanisms (similar to digital panopticism [11], guiding students to self-regulate), classification mechanisms (algorithms labeling students based on behavioral data [12]), and evaluation mechanisms (AI scoring systems [13] redefining academic evaluation criteria). The introduction of

Foucault's theory enables this framework to move beyond superficial discussions of "efficiency improvement" and deeply reveal how AI, as a new source of power, participates in the redefinition of knowledge, authority, and qualified subjects in education. It clarifies how rational individual technological choices converge into structural power shifts, thereby demonstrating the deep logic and fundamental impact of power changes.

In summary, the integration of the Technology Acceptance Model and Foucault's theory forms the theoretical core of this framework. Reliance solely on the Technology Acceptance Model may limit analysis to superficial narratives, failing to explain the far-reaching social consequences of technology application; exclusive use of Foucault's theory may render criticism abstract when divorced from empirical starting points. The combination of the two enables capturing micro-level power flows from supervisors' and students' specific use of intelligent technology, while understanding the structural significance of these behaviors at the level of macro power structure changes, thereby constructing a complete explanatory chain from the micro to the macro and from initiation to deepening.

2.2. Structural design: structured interpretation across operational, supervisory, and institutional levels

Building on the integrated framework, this paper further constructs the functional stratification-power checks and balances three-dimensional analytical model (Figure 1). It should be noted that this model does not divide reality into three independently operating physical layers but provides three interwoven and dynamically progressive theoretical perspectives to reveal the complete logic of power changes in the collaborative system between AI and human supervisors. Specifically, the operational level serves as the micro empirical starting point, focusing on the efficiency empowerment function of AI in performing specific academic tasks such as literature reviews and data processing. Its core dynamics stem from the "perceived usefulness" and "perceived ease of use" identified by the Technology Acceptance Model, forming a dynamic balance of efficiency empowerment/power transfer, where power transfer constitutes the micro behavioral foundation of power flow. The supervisory level, as the meso-level conflict field, shifts the analytical focus to human supervisors' supervisory functions in guiding academic directions, safeguarding academic ethics, and evaluating innovation. At this level, the traditional authority of supervisors engages in direct struggle with the emerging algorithmic authority of AI, with core dynamics rooted in the persistent tension between algorithmic objectivity and human judgment, which can be examined through Foucault's Power-Knowledge Theory. The institutional level, as the macro institutional framework, focuses on the macro regulation of AI use through policies, norms, and guidelines formulated by state and university institutions. Its core dynamics lie in responding to technological shocks to uphold educational purposes and cultural traditions, reflecting institutional efforts of various countries to preserve core educational concepts amid technological change, with its power consolidation pattern interpretable from an institutionalist theoretical perspective.

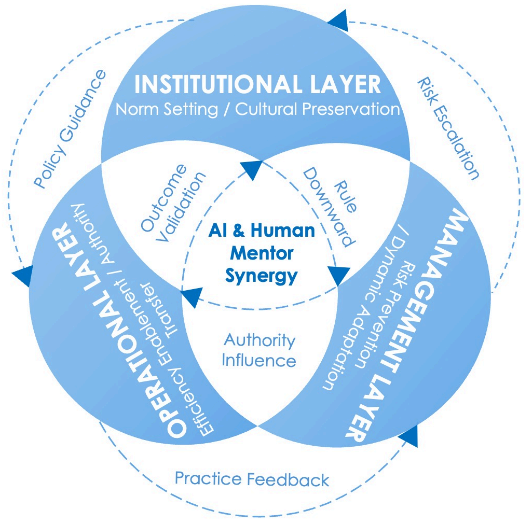

Figure 1. Functional stratification-power checks and balances three-dimensional analytical model

From the perspective of internal logical correlation, the three dimensions are not linearly parallel but present a dynamic relationship of both progressive generation and structural embeddedness. The operational level, driven by voluntary power transfer by supervisors and students based on efficiency considerations, forms the empirical starting point of this framework; once these scattered individual transfer behaviors converge in core links of academic judgment and ethical regulation, conflicts inevitably arise, thereby giving rise to the need for power struggle at the supervisory level—the traditional authority of human supervisors and the algorithmic authority of AI contend over the right to define knowledge and evaluate subjects, which constitutes the meso-level process of power conflict and renegotiation. Ambiguous norms and accountability vacuums emerging from the struggle further drive state and university institutions to respond through institutional design, solidifying the struggle outcomes into new power boundaries and accountability divisions, thus forming a pattern of power consolidation at the institutional level. Meanwhile, the institutional level is not merely a passive response terminal; as the "underlying rules", it is embedded in the operational and supervisory levels in reverse, predefining the behavioral boundaries of supervisors' and students' technology use and the legitimate space for supervisors' supervisory struggles, thereby forming a complete cycle that progresses from micro behaviors to macro institutions and shapes micro practices from top to bottom. Through this progressive explanatory chain of operational transfer-supervisory struggle-institutional consolidation, this framework (Table 1) breaks the static pattern of simply juxtaposing the three dimensions, systematically revealing the entire process of power balance relations from emergence, conflict to institutionalization in graduate education in the AI era, and providing a theoretical tool with both structural characteristics and dynamic explanatory power for understanding and analyzing this complex phenomenon.

Table 1. Functional stratification-power checks and balances analytical framework

Analytical Dimension	Logical Positioning	Core Function	Core Dynamics	Power Dynamics	Dominant Theoretical Perspective
Operational Level	Micro empirical starting point: behavioral foundation of power flow, providing initial dynamics for subsequent struggles	Efficiency empowerment: performing standardized and procedural tasks (literature retrieval, grammar checking, code generation, etc.)	Perceived usefulness and perceived ease of use	Power transfer: Supervisors and students voluntarily transfer cognitive and operational power to AI in pursuit of efficiency, forming micro behavioral dynamics	Technology Acceptance Model (TAM)
Supervisory Level	Meso-level conflict field: core arena of power struggle, convergence point of transfer behaviors	Ethical regulation and value guidance: guiding academic directions, evaluating originality, cultivating scientific spirit, and providing emotional support	Algorithmic objectivity vs. human judgment	Power struggle: The traditional authority of human supervisors and the "algorithmic authority" of AI engage in conflict and renegotiation over the right to define knowledge and evaluate subjects	Foucault's Power-Knowledge Theory
Institutional Level	Macro institutional framework: top-level design for power consolidation, embedded in the first two dimensions as underlying rules in reverse	Normative institutionalization and cultural maintenance: formulating policies and guidelines, adjusting organizational structures and accountability systems	Responding to technological shocks to uphold educational purposes and cultural traditions	Power consolidation: Institutions confirm new power boundaries and accountability divisions through institutional design, embed cultural values into technology application, and shape operational and supervisory spaces in reverse	Institutionalist Theory

The core value of this three-dimensional analytical model lies in its structured explanatory power. It not only clearly distinguishes the roles of AI at different levels but also reveals the differentiated logic of power operation at each level. First, at the operational level, the model focuses on the specific dynamic of power "transfer", clarifying that such transfer is not passive deprivation but an active choice based on the "perceived usefulness" identified by the Technology Acceptance Model (TAM), thereby providing an empirical foundation for understanding the micro starting point of power transfer. Further, at the supervisory level, the model shifts to analyzing the complex process of power "struggle", identifying the core conflict as the contention between human supervisors' "traditional authority" and AI's "algorithmic authority", and drawing on Foucault's theory to deeply analyze how this struggle revolves around the right to define knowledge and evaluate subjects. Finally, at the institutional level, the model focuses on the form of power "consolidation", adopting an institutionalist perspective to explain how state and university institutions respond to power vacuums or conflicts caused by technology through policies, regulations, and guidelines, and attempt to embed local educational concepts and cultural traditions (e.g., humanism, national interests) into the technology-institution structure, thereby shaping the long-term pattern of power relations. Through the progressive correlation of the three levels, the framework systematically and comprehensively addresses the core research questions raised, providing a structured theoretical tool for understanding and analyzing power balance relations in graduate education in the AI era.

3. International comparison: in-depth analysis of cases in China, the United States, and Russia

Based on the functional stratification-power checks and balances analytical framework, this paper selects typical practices of AI-supervisor collaboration in graduate education in the United States, Russia, and China for comparative research. The higher education systems of the three countries exhibit significant differences in governance structure, academic traditions, and state-market relations, and this institutional diversity covers the main types of governance logics, making it an ideal field for exploring how institutional environments shape human-AI collaboration paths. From a three-dimensional hierarchical perspective, the framework systematically deconstructs the implicit systems of power transfer, authority reconstruction, and institutional adaptation in the practices of the three countries, not only providing an empirical analytical path with national comparative significance for understanding the complex dynamics of human-AI collaboration in graduate education in the AI era but also testing the applicability and explanatory power of this theoretical framework in explaining educational changes in different institutional contexts.

3.1. The United States: a "functional outsourcing" model driven by the market

Taking top private research universities such as Stanford University as typical representatives, the U.S. graduate education system is characterized by a weak supervisor responsibility system and a free-market culture, tending to position AI as an academic service tool that supervisors and students can independently choose. Its collaborative system exhibits clear dynamic characteristics at the operational, supervisory, and institutional levels.

At the operational level, the most prominent feature is extensive and voluntary power transfer. Practices at Stanford University and other institutions show that the application of tools such as ChatGPT-4.0, Scite, and Elicit is mainly driven by individual needs of supervisors and students. An anonymous survey of doctoral students at Ivy League universities in 2024 revealed that over 60% of respondents use AI to assist with tasks such as literature retrieval, programming, and data visualization, with the core motivation being "significant time savings" and "access to more inspiration" [14]. Such voluntary adoption behaviors driven by the perceived usefulness in the Technology Acceptance Model efficiently "outsource" basic and procedural academic work to AI systems, freeing supervisors from detailed supervision tasks and enabling them to focus on academically leading work of higher core value.

At the supervisory level, supervisors' authority shows a trend of managerial transformation, while algorithmic authority is strictly limited to the tool level. Under the framework of the weak supervisor responsibility system, supervisors act more as "academic advisors", with their core authority manifested in directional guidance, provision of academic network resources, and final quality control of outcomes, rather than full-process control of academic activities. The intervention of AI has not shaken supervisors' core authority but is regarded as an effective tool to improve team efficiency, with power struggles increasingly focusing on academic integrity. For example, Stanford University issued the *Generative AI Policy Guidance* on February 16, 2023, explicitly stipulating that content generated by AI used by students must be clearly disclosed and prohibited for assignments or examinations, and teachers will use detection software to review AI usage [15]. This measure essentially confines algorithmic authority to the category of auxiliary tools and transforms supervisors' supervisory role into a "managerial" function of overseeing students' compliant use of tools.

At the institutional level, the United States has formed a decentralized system of policy patches, reflecting a deep commitment to liberal culture. Institutional responses in U.S. universities are distinctly decentralized, lacking uniform national mandatory regulations; instead, individual universities, departments, and even courses independently formulate "policy patches" [16]. These policies are generally centered on the principles of transparency, accountability, and honesty, reflecting full respect for individual choice and academic freedom, and essentially represent a sustained commitment to the core liberal and market-oriented culture of U.S. higher education. The logic of this institutional design lies in the belief that a market-driven academic ecosystem can achieve self-regulation through clear rules and severe punishment for violators, thereby maximizing the benefits of AI while effectively controlling its risks.

In summary, the relationship between AI policies and supervisors' authority in the United States can be summarized as a market-driven functional outsourcing model, with its core logic as follows: at the operational level, driven by perceived usefulness, supervisors and students voluntarily outsource basic academic work such as literature retrieval and programming to AI tools, achieving efficiency improvement and liberation of supervisors' roles; at the supervisory level, under the weak supervisor responsibility system, supervisors' authority transforms into a "managerial" role in directional guidance and quality control, algorithmic authority is strictly limited to auxiliary tools, and power struggles focus on academic integrity; at the institutional level, a decentralized system of policy patches has taken shape, educational institutions independently formulate rules, uphold liberal and market-oriented cultural traditions, and promote self-regulation of the academic ecosystem through clear norms and punishment mechanisms. Essentially, this model promotes the natural penetration of AI technology through market systems, constructing a dynamic balance mechanism covering power transfer, authority reconstruction, and institutional adaptation while safeguarding academic freedom.

3.2. Russia: an "auxiliary integration" model led by the state

Taking practices at Herzen State Pedagogical University in training future teachers as a representative case, Russia's graduate education system is deeply shaped by national planning and collectivist traditions, tending to systematically position AI as a standardized tool serving national educational strategies and assisting and strengthening the existing human supervisor system. Its collaborative system exhibits unique dynamic characteristics at the operational, supervisory, and institutional levels.

At the operational level, power transfer is notably limited and controlled. Unlike the unregulated diffusion of market-driven tools, AI applications in Russian education highlight top-level planning. For example, the adaptive Electronic Information Educational Environment (EIEE) platform developed by Herzen State Pedagogical University integrates AI chatbots [17], explicitly limiting their functions to handling repetitive and standardized communication tasks such as course schedule inquiries, learning resource recommendations, and assignment deadline reminders. This design ensures that supervisors and students only transfer low-level cognitive power related to "information transmission" and "process management", with core teaching supervision activities remaining under the control of human supervisors. Power transfer is entirely geared toward efficiency improvement rather than role replacement.

At the supervisory level, supervisors' authority undergoes an empowering transformation, while AI is clearly instrumentalized. AI chatbots undertake pre-stage standardized interactions, freeing human supervisors from routine tasks to focus on higher-level teaching activities such as heuristic teaching, personalized guidance, and complex problem-solving. The core goal of this project is to train future teachers to "use AI to solve professional tasks" [17], which presupposes that the output efficiency of AI must be supervised, evaluated, and reflected upon by human supervisors. In this framework, algorithmic authority is thoroughly instrumentalized and does not constitute an independent source of authority; its value depends entirely on supervisors' judgments based on teaching objectives and effects. Rather than weakening supervisors' authority, technology further consolidates their dominant position in core teaching activities through empowerment.

At the institutional level, Russia has formed a strategic capacity-building system, reflecting a deep commitment to collectivist culture. This practice responds to Russia's national strategy for a digital educational environment, with goals extending beyond mere tool introduction to systematically developing new courses to cultivate future teachers' professional capabilities in human-AI collaboration, aligning institutionally with the Unified AI Technology Training Program for Russian Universities promoted by the Ministry of Science and Higher Education of the Russian Federation [18]. This top-down capacity-building approach profoundly demonstrates Russia's long-standing commitment to state leadership, collectivism, and technological sovereignty: AI applications are positioned as an integral part of the modernization of the national education system, with their development paths and boundaries uniformly planned by national strategies to ensure alignment with national interests and educational philosophies.

In summary, the relationship between AI policies and supervisors' authority in Russia can be summarized as a state-led auxiliary integration model, whose essence lies in positioning AI as a standardized tool serving national educational strategies and achieving human-AI collaboration through limited and controlled power transfer. At the operational level, AI only undertakes repetitive and standardized information transmission and process management tasks such as course inquiries, resource recommendations, and assignment reminders, with core teaching supervision power remaining under the control of human supervisors; at the supervisory level, AI is thoroughly instrumentalized, algorithmic authority does not constitute an independent source of authority, and its output efficiency must be evaluated and reflected upon by supervisors, with technological empowerment instead strengthening supervisors' dominant position in higher-level teaching activities; at the institutional level, from the EIEE platform at Herzen State Pedagogical University to the national unified training program, all reflect Russia's commitment to state leadership, collectivism, and technological sovereignty. AI applications are integrated into the overall process of education system modernization, with their development paths and boundaries uniformly planned by national strategies to ensure alignment with national interests and educational philosophies.

3.3. China: a "power struggle" model under institutional constraints

Represented by "Double First-Class" universities such as Tsinghua University, China's graduate education system is characterized by the institutional rigidity of the Supervisor as Primary Responsible Party system, with its deep cultural logic shaped by the Confucian tradition of respecting teachers and valuing education and collectivist accountability ethics. This dual institutional-cultural framework creates persistent tension between AI penetration and existing normative frameworks, with its collaborative system exhibiting particularly complex dynamic characteristics at the operational, supervisory, and institutional levels.

At the operational level, driven by the global technological wave and active deployment by domestic technology enterprises, domestic large-scale models (e.g., DeepSeek-V4, Ernie Bot 5.0, Tongyi Qwen 3.6-Max) have rapidly gained popularity in universities with their powerful capabilities in code generation, data analysis, Chinese-English translation, and policy text analysis, presenting explosive and comprehensive power transfer. Empirical data show that as high as 96.8% of Chinese university students actually use AI for academic tasks, and their usage behavior is significantly driven by perceived usefulness and ease of use [19], with the speed and scope of power transfer even surpassing those in the United States. Basic and procedural academic work has been extensively transferred to AI systems, and students' reliance on AI has extended from auxiliary tools to core research links.

At the supervisory level, the rapid power transfer at the operational level immediately triggers severe conflicts with the institutional requirements of the Supervisor as Primary Responsible Party system. On the one hand, the system stipulates that supervisors bear near-unlimited responsibility for graduate students' academic achievements, academic conduct, and even graduation prospects; on the other hand, students' research processes are increasingly detached from supervisors' oversight due to deep reliance on AI, with weekly reports, codes, or draft theses submitted possibly mostly generated by AI, making it difficult for supervisors to identify students' true abilities and original contributions, and the supervisory role faces the risk of hollowed-out authority. Meanwhile, algorithmic authority implicitly rises with technological advantages: current AI educational applications generally lack educational theoretical support and teaching guidance, leaving students highly susceptible to passive reliance on algorithmic authority [20]. At this point, algorithmic authority is no longer merely a tool or subordinate but a powerful implicit collaborator and even competitor behind students, directly challenging supervisors' traditional authority in knowledge, methods, and experience. This is the core conflict of China's model, triggering deep dilemmas of authority erosion, trust crisis, and responsibility-power mismatch.

At the institutional level, facing "out-of-control" operations and "chaotic" supervision, institutional responses exhibit a lagging feature of normative catch-up. On September 10, 2024, the Chinese Academy of Sciences issued the Integrity Reminder on Standardizing the *Use of Artificial Intelligence Technology in Scientific Research Activities* [21], followed by many universities issuing documents such as *Notice on Standardizing the Use of AI Tools in Undergraduate Graduation Projects (Theses)* in the second half of 2024, generally drawing "red lines" for AI use, such as prohibiting the direct generation of core thesis content, and reaffirming supervisors' supervisory responsibilities. Behind these institutional efforts lies a deep commitment to China's traditional Confucian culture of respecting teachers and valuing education and collectivist accountability ethics, attempting to reintegrate technological disorder into the existing ethical framework emphasizing teacher dignity and absolute supervisor responsibility through a series of policy patches, though their effectiveness remains severely challenged.

In summary, the relationship between AI policies and supervisors' authority in China can be summarized as a power struggle model under institutional constraints, with its core contradiction lying in the intense conflict between market-driven rapid technological penetration at the operational level and the institutional rigidity of the Supervisor as Primary Responsible Party system at the institutional level: explosive power transfer driven by high perceived usefulness at the operational level, hollowed-out supervisor authority and implicitly rising algorithmic authority coexisting at the supervisory level, and lagging normative catch-up and commitment to traditional Confucian culture at the institutional level. Essentially, this model is a dynamic game process between technological penetration and institutional constraints, highlighting the unique tensions faced by developing countries in AI educational applications.

4. Deconstruction and reconstruction: building a China-style three-tier collaborative framework

4.1. Case comparison: a comparative analysis of AI collaborative systems in the three countries

By constructing the functional stratification-power checks and balances analytical framework, this study conducts an in-depth comparison of human-AI collaborative systems in graduate education in China, the United States, and Russia. The framework decomposes educational scenarios into three functional dimensions—operational, supervisory, and institutional—and systematically deconstructs the dynamic processes of technology embedding in education under different institutional and cultural backgrounds along three analytical threads: power transfer paths, authority reconstruction logic, and institutional adaptation systems. The results reveal that due to differences in institutional endowments and cultural traditions, the three countries have formed distinct collaborative models. The United States has developed a market-driven functional outsourcing model, where AI is incorporated into supervisors' project management functions as a market-oriented technological tool, with power transfer characterized by voluntary authorization by students and selective acceptance by supervisors, and authority reconstruction mainly reflected in the supplementation of professional authority by instrumental rationality. Russia presents a state-led auxiliary integration model, where AI strengthens supervisors' supervisory capabilities through state-led technological empowerment, with power transfer manifested as top-down technological implantation, and authority reconstruction highlighting the absolute dependence of technological authority on supervisors' authority. China, however, features a power struggle model under institutional constraints, with profound tensions between technological abuse at the operational level and institutional rigidity at the institutional level. Power transfer is in a contradictory state of de facto substitution and institutional constraint, and authority reconstruction is concentrated in the implicit challenge of algorithmic authority to traditional supervisor authority. This algorithmic authority does not stem from legal or traditional authorization but is a cognitive authority generated based on technological characteristics, with its legitimacy rooted in three psychological systems: perceived data objectivity, perceived efficiency advantage, and perceived breadth of knowledge, and it presents differentiated manifestations in different institutional and cultural contexts: in the United States, it becomes implicit authority in market-driven instrumental authorization; in Russia, it resides in auxiliary authority as a technological appendage under state leadership; in China, it highlights implicit competitive authority in institutional struggles under institutional constraints. The conceptual refinement and framework construction above not only provide a theoretical lens for understanding cross-cultural reconstruction of educational power in the digital age but also offer practical implications for building a human-AI collaborative new educational ecosystem by revealing deep tensions between technological logic and institutional logic—that is, to achieve a dynamic balance between technological rationality and educational ethics while safeguarding supervisors' core educational functions and reasonably releasing the technological efficacy of AI.

4.2. Strategic responses: building a hierarchically linked three-tier collaborative framework

Confronting the real dilemmas under China's Supervisor as Primary Responsible Party system, this study attempts to transcend the binary thinking of "banning" and "laissez-faire", constructing a layered, dynamically linked, and actionable three-tier collaborative framework to achieve a balance between technological empowerment and institutional adherence.

At the operational level, focus on the institutional construction of a positive and negative list of human-AI functions, with the underlying intention of establishing structural principles for human-AI function allocation rather than merely providing a behavioral guide. To this end, the negative list and positive list are assigned different institutional functions: the former, as a bottom-line provision, explicitly prohibits the use of AI to generate core arguments, research designs, raw data, and conclusions of degree theses, thereby defining the minimum contribution boundary of "humans" in degree awarding, with violations held accountable in accordance with academic misconduct norms; the latter, as an empowerment provision, encourages and guides students to use AI for non-original thinking tasks such as initial literature screening, language polishing, code debugging, format typesetting, and data visualization, with the development of relevant skills integrated into graduate academic norms and information literacy courses [22]. On this basis, an academic AI assistant integrated into the campus network can be developed, with functions preset in accordance with the positive and negative lists and automatic logging of usage records to achieve traceability [23, 24], thereby shifting ethical governance from post-event accountability to in-process embedding.

At the supervisory level, promote the transformation of supervisors' role into process supervisors, with the key not only being capacity improvement but also the institutional reconfirmation of the authority required for this role. The crisis of hollowed-out supervisor authority stems from the dilution of knowledge authority by technology and the failure of institutional authority to keep pace, so the reconstruction of the supervisory framework must be approached from two paths. On the one hand, shift the focus of supervisors' supervision from result review to process supervision, requiring students to submit research logs containing AI interaction records on a regular basis and replicate key analytical processes in group meetings to ensure the authenticity and originality of research processes; meanwhile, systematically train supervisors in AI literacy, enabling them to understand the capabilities, potential risks, and effective detection methods of mainstream AI tools, and enhancing their supervisory competence in the digital age [25]. On the other hand, explicitly incorporate the above process supervision functions into supervisor responsibility regulations, granting supervisors procedural powers such as accessing students' AI usage logs and requiring replication of analytical processes, forming institutionalized procedural supervisory power, and clarifying boundaries with student privacy protection. University-provided AI originality detection tools for supervisors should maintain their positioning as auxiliary judgment rather than independent adjudication—final professional judgments must be made by supervisors based on a comprehensive understanding of students' research processes, preventing algorithmic detection from encroaching on supervisors' decision-making space.

At the institutional level, explore the dynamic adjustment of a supervisor rights and responsibilities list, with the core aim of alleviating the contradiction of power-responsibility mismatch from the institutional root and introducing iterative governance experimentation thinking rather than one-time revision. Specifically, select some universities or departments to pilot a flexible supervisor responsibility system, test different degrees of rights and responsibilities relaxation schemes, and establish supporting process assessment and risk early warning frameworks. In adjusting the connotation of rights and responsibilities, clarify that supervisors bear primary responsibility for students' academic integrity and core competence cultivation, but under the premise of compliant AI use, they may be exempted from unlimited joint liability for minor errors; establish public support positions such as AI technical consultants or data scientists at the department level, forming a new collaborative education model of supervisor-led, expert-assisted, AI-empowered [26]. Meanwhile, establish a flexible evaluation system, introducing three process-based evaluation indicators that AI is least likely to simulate—problem-posing ability, critical dialogue ability, and error correction ability in research processes—in graduate academic assessments and supervisor evaluations, serving as sustained testing tools to measure the operational effectiveness of the collaborative framework, guiding supervisors and students back to the essence of education and innovation through evaluation orientation.

The above three-tier framework achieves dynamic linkage through boundary regulation and ethical embedding at the operational level, role transformation and institutionalization of authority at the supervisory level, and adjustment of rights and responsibilities and iterative governance at the institutional level, not only responding to the practical needs of graduate training in the AI era but also upholding the ethical bottom line of the Supervisor as Primary Responsible Party system, thereby moving the deep reconciliation between technological rationality and educational rationality from concepts to institutional schemes.

4.3. Research prospects: power reconstruction and adherence to educational missions in institutional innovation

As an exploratory achievement in 2025, although this study has initially constructed the functional stratification-power checks and balances analytical framework and revealed multi-dimensional tensions in AI-supervisor collaborative system, the characterization of cases in China, the United States, and Russia still relies on more first-hand empirical data for deepening due to the limited availability of public data. For example, at the operational level, specific cases of implementing the "positive and negative list of human-AI functions" in individual universities can be supplemented to analyze its actual impact on students' academic autonomy and supervisors' supervisory effectiveness; at the supervisory level, in-depth interview data can be introduced to reveal cognitive conflicts and capacity challenges faced by supervisors in the process of role transformation; at the institutional level, longitudinal tracking research can be conducted to evaluate the long-term effects of the "dynamic adjustment of the supervisor rights and responsibilities list" on the educational ecosystem. On this basis, future research can be expanded in three directions along this framework. First, conduct large-scale questionnaire surveys to quantify the perceived differences of algorithmic authority among graduate students of different disciplines and grades, and verify psychological motivations and institutional constraints in its formation system combined with structural equation modeling. Second, use in-depth interviews and participant observation to capture the dynamic changes in supervisor-student interaction patterns from tool dependence to collaborative innovation, and further extract universally applicable collaborative strategies. Third, adopt field experimental design to test the implementation effects of the three-tier collaborative framework under controlled variables, for example, comparing the actual impact of different "positive and negative list" designs on students' academic integrity and innovation capabilities.

5. Conclusion

In short, AI technology brings not only superficial challenges of efficiency improvement but also profound reconstruction of educational power structures and knowledge production logic. Only through proactive institutional innovation amid persistent tensions between technological rationality and educational ethics—such as constructing a legitimacy evaluation system for algorithmic authority and designing a feedback regulation framework for triangular interaction among supervisors, AI, and students—can we ensure that technology truly serves the fundamental mission of graduate education: cultivating high-level talents with original spirit and holistic competence, thereby achieving a paradigm leap from technology adaptation to technology empowerment. The above research extensions not only deepen reflections on limitations but also strengthen the empirical foundation of theoretical construction through the layout of specific research paths, reflecting the norms of academic language and the inherent logic of research.

References

[1] Choung, H., David, P., & Ross, A. (2023). Trust in AI and its role in the acceptance of AI technologies. *International Journal of Human–Computer Interaction, 39*(9), 1727-1739.

[2] Ali, I., Warraich, N. F., & Butt, K. (2025). Acceptance and use of artificial intelligence and AI-based applications in education: A meta-analysis and future direction. *Information Development, 41*(3), 859-874.

[3] Marjerison, R. K., Dong, H., Kim, J.-M., Zheng, H., Zhang, Y., & Kuan, G. (2025). Understanding User Acceptance of AI-Driven Chatbots in China's E-Commerce: The Roles of Perceived Authenticity, Usefulness, and Risk. *Systems, 13*(2), 71. https: //doi.org/10.3390/systems13020071

[4] Chiu, M. L. (2025). Exploring user awareness and perceived usefulness of generative AI in higher education: The moderating role of trust. *Education and Information Technologies*, 1-35.

[5] Selwyn, N. (2022). The future of AI and education: Some cautionary notes. *European Journal of Education, 57*(4), 620-631.

[6] Madaio, M., Blodgett, S. L., Mayfield, E., & Dixon-Román, E. (2022). Beyond "fairness": Structural (in) justice lenses on ai for education. In The ethics of artificial intelligence in education (pp. 203-239). *Routledge*.

[7] Bareis, J., & Katzenbach, C. (2022). Talking AI into being: The narratives and imaginaries of national AI strategies and their performative politics. *Science, Technology, & Human Values, 47*(5), 855-881.

[8] Davis, F. D. (1989). *Technology acceptance model: TAM. In MN Al-Suqri & AS Al-Aufi (Eds.), Information Seeking Behavior and Technology Adoption* (pp. 5-219).

[9] Dai, Y., Lai, S., Lim, C. P., & Liu, A. (2023). ChatGPT and its impact on research supervision: Insights from Australian postgraduate research students. *Australasian Journal of Educational Technology, 39*(4), 74-88.

[10] Foucault, M. (2020). Power/knowledge. In The new social theory reader (pp. 73-79). *Routledge*.

[11] Zhang, J., & Chen, M. (2021). Returning to the Value of Education: A Path to Resolve the Hidden Worries of Educational Discipline Under Digital Panopticism. *Modern Distance Education Research, 33*(4), 33-42. DOI: CNKI: SUN: XDYC.0.2021-04-005.

[12] Angelides, M.C. (1992). *Developing the didactic operations for intelligent tutoring systems: A synthesis of artificial intelligence and hypertext.*

[13] McConvey, K., Guha, S., & Kuzminykh, A. (2023, April). *A human-centered review of algorithms in decision-making in higher education*. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (pp. 1-15).

[14] Lyu, W., Wang, Y., Sun, Y., & Zhang, Y. (2025, July). *Will Your Next Pair Programming Partner Be Human? An Empirical Evaluation of Generative AI as a Collaborative Teammate in a Semester-Long Classroom Setting.* In Proceedings of the Twelfth ACM Conference on Learning@ Scale (pp. 83-94).

[15] Stanford Office of Community Standards. Generative AI Policy Guidance. (2023-02-16). Retrieved October 27, 2025, from https: //communitystandards.stanford.edu/generative-ai-policy-guidance

[16] Vats, V., Binta Nizam, M., Liu, M., Wang, Z., Ho, R., Sai Prasad, M., ... & Davis, J. (2024). *A survey on human-ai teaming with large pre-trained models.* arXiv e-prints, arXiv-2403.

[17] Vlasova, E. Z., Avksentieva, E. Y., Goncharova, S. V., & Aksyutin, P. A. (2019). Artificial intelligence-The space for the new possibilities to train teachers. *Espacios, 40*(9), 17.

[18] Zmyzgova, T. R., Polyakova, E. N., & Karpov, E. K. (2020, May). *Digital transformation of education and artificial intelligence.* In 2nd International Scientific and Practical Conference "Modern Management Trends and the Digital Economy: from Regional Development to Global Economic Growth"(MTDE 2020) (pp. 824-829). Atlantis Press.

[19] Li, K. (2023). Determinants of college students' actual use of AI-based systems: An extension of the technology acceptance model. *Sustainability, 15*(6), 5221.

[20] Ouyang, F., Zheng, L., & Jiao, P. (2022). Artificial intelligence in online higher education: A systematic review of empirical research from 2011 to 2020. *Education and Information Technologies, 27*(6), 7893-7925.

[21] Chinese Academy of Sciences. Integrity Reminder on Standardizing the Use of Artificial Intelligence Technology in Scientific Research Activities. (2024-09-11). Retrieved October 28, 2025, from https://www.cas.cn/sygz/202409/t20240910_5031186.shtml

[22] Wu, F. T., Yang, C. Y., & Li, T. (2024). Human-AI Collaborative Design in Smart Learning Environments. *E-education Research, 45*(2), 84-90. DOI: 10.13811/j.cnki.eer.2024.02.012.

[23] Lai, C. (2023, June). *Establishment of Educational AI System Based on Man-machine Coordination.* In ITEI 2022: Proceedings of the 2nd International Conference on Internet Technology and Educational Informatization, ITEI 2022, December 23-25, 2022, Harbin, China (p. 479). European Alliance for Innovation.

[24] Zhang, J. (2021). Computer assisted instruction system under artificial intelligence technology. *International Journal of Emerging Technologies in Learning (iJET), 16*(5), 4-16.

[25] Wu, W., Burdina, G., & Gura, A. (2023). Use of artificial intelligence in teacher training. *International Journal of Web-Based Learning and Teaching Technologies (IJWLTT), 18*(1), 1-15.

[26] Yu, J. (2024). Research on application and quality evaluation of intelligent teaching model of AI three teachers in the course of motion control technology. *International Journal of Education and Humanities*, *15*(3), 134-139.