Artificial Intelligence Education Studies
Volume 1, Issue 3, 2025, Page 70–82
https://doi.org/10.6914/aiese.010306

Artificial Intelligence Education Studies

Productive Entanglements: A Quantum-Agential Framework for Cognition, Agency, and Ethics in AI-Mediated Education

Gang DENG
Qilu Institute of Technology

Abstract

This paper introduces the Quantum-Agential framework to address the "foundational gap" in AI-mediated education, where classical dualisms fail to capture the entangled realities of human machine learning. Drawing on Karen Barad's Agential Realism and quantum-like models of cognition, the framework reconceptualizes cognition, agency, and ethics as emergent properties of material-discursive apparatuses rather than attributes of isolated individuals or tools. By dissolving binaries such as human/machine and subject/object, it provides a relational ontology and a formal language for modeling uncertainty, contextuality, and entanglement in learning processes. Case studies of co-writing with generative AI and adaptive tutoring systems illustrate how this approach reframes cognitive load, distributed agency, and systemic ethics. The framework advances post-humanist pedagogy, emphasizes educator roles as architects of learning apparatuses, and proposes diffractive methodologies for research. It offers a principled path toward ethically responsible, cognitively aware, and human-centered futures for AI in education.

Keywords AI in Education; Agential Realism; Quantum-like Cognition; Distributed Agency; Ethics

1 Introduction: The Foundational Gap in an Age of Intelligent Machines

The rapid and pervasive integration of Artificial Intelligence into Education (AIEd) has precipitated a moment of profound transformation and considerable uncertainty. Technologies ranging from adaptive learning systems that tailor educational experiences to individual student needs, to sophisticated generative models like ChatGPT capable of producing human-like text, are reshaping pedagogical practices, automating assessments, and personalizing learning pathways on an unprecedented scale. This technological proliferation, however, has outpaced the deliberate development of robust pedagogical, ethical, and theoretical frameworks to guide its use, creating what can be termed a "foundational gap". The current discourse surrounding AIEd often reflects this gap, oscillating between the uncritical optimism of technological utopianism and the profound anxieties of dystopian futures marked by cognitive degradation and ethical collapse. This polarization signals a collective lack of a stable, integrated, and nuanced understanding of AI's capabilities and limitations within the complex ecosystem of education.

Evidence suggests that educational institutions are adopting a largely reactive posture, implementing powerful technologies before their deeper, long-term implications are fully comprehended or addressed by well-conceived guiding principles. This reactive stance means that current AIEd practices are often constructed on unstable conceptual ground, risking unforeseen and undesirable consequences for learners, educators, and the very mission of education itself. The problem, however, is not merely a matter of speed—that technology is simply evolving too quickly for pedagogy and policy to keep pace. Rather, this foundational gap is symptomatic of a more fundamental ontological mismatch. We are attempting to understand and govern the phenomena of modern AIEd using classical conceptual tools—grounded in dualisms of subject/object, human/machine, mind/world, and knower/known—that are increasingly inadequate for capturing the relational, entangled, and co-constitutive reality of AI-mediated learning.

 Article History Received: June 1, 2025; Accpted: July 12, 2025; Published: September 30, 2025. ***Artificial Intelligence Education Studies***, ISSN 3079-8086 (print), ISSN 3079-8094 (online), a Quarterly, founded on 2025, published by Creative Publishing Co., Limited. Email: wtocom@gmail.com, https://ai-es.org, https://cpcl.hk.

This mismatch becomes starkly evident in common educational scenarios. When a student co-creates an essay with a generative AI, the classical question "Who is the author?" becomes profoundly difficult to answer. The locus of knowing is no longer a singular, clearly bounded entity but a hybrid, a dynamic coupling of human and machine cognition. Similarly, when an adaptive tutoring system guides a learner's path, the agency in the learning process is not located solely in the student or the machine but is distributed and emergent. Our existing theoretical frameworks, which presuppose the existence of separate, pre-defined entities that subsequently "interact," struggle to account for these new realities. The persistent and unproductive oscillation between utopian and dystopian narratives is a direct symptom of this conceptual inadequacy. By attempting to force a non-dualistic phenomenon—the deep entanglement of human and artificial intelligence—into a dualistic framework that pits human against machine, the discourse is unable to forge a productive path forward. The ontological mismatch is the true source of the foundational gap, and closing it requires more than just updated policies; it requires a fundamental shift in our understanding of the very nature of learning, knowing, and being in an age of intelligent machines.

This paper addresses this challenge by posing the following research question: How can a post-humanist, relational ontology reframe our understanding of cognition, agency, and ethics in AI-mediated learning ecosystems, and what are the implications for pedagogical design and future inquiry? To answer this, it develops and proposes a novel theoretical model: the Quantum-Agential framework. This framework synthesizes the radical onto-epistemology of Karen Barad's Agential Realism with the formal and conceptual tools of Quantum-like models of cognition. It argues that this integrated approach provides a more adequate ontological and analytical apparatus for navigating the complexities of AIEd. Agential Realism offers a post-humanist worldview that dissolves classical dualisms, while quantum-like models provide a formal language to describe the non-deterministic, contextual, and entangled dynamics that characterize cognition and interaction within these new learning assemblages.

By providing a new ontological foundation, the Quantum-Agential framework aims to move the discourse beyond the simplistic binary of utopianism versus dystopianism. It offers a way to analyze AIEd phenomena not as interactions between separate entities but as emergent properties of a complex, entangled system. This paper will first construct this theoretical framework in detail. It will then demonstrate its analytical power by applying it to reinterpret core issues in AIEd, including cognition, agency, and ethics, using illustrative case studies. Finally, it will explore the profound implications of this framework for pedagogical design, the evolving role of the educator, and the future of educational research. The ultimate goal is to offer a robust conceptual lens through which to build a more principled, cognitively-aware, and ethically-responsive future for AI in education.

To situate the proposed Quantum-Agential framework and underscore its unique contribution, it is useful to contrast it with dominant theoretical paradigms that have historically informed AIEd. The following table provides a comparative analysis across key dimensions, illustrating the conceptual shift this paper advocates.

This comparison highlights how the Quantum-Agential framework moves beyond both individualistic and purely social models to offer a radically relational and post-humanist perspective. By taking the entire apparatus as the unit of analysis, it provides a more holistic and systemic way to understand the complex dynamics of modern AIEd, where the boundaries between human, machine, and context are increasingly blurred.

2 Forging the Theoretical Apparatus: A Quantum-Agential Framework for AIEd

To address the ontological mismatch at the heart of the foundational gap, a new theoretical apparatus is required—one that can account for the relational, entangled, and emergent nature of AI-mediated learning. This section constructs such an apparatus by synthesizing two powerful, post-classical theoretical streams: Karen Barad's Agential Realism, which provides the ontological grounding, and Quantum-like models of cognition, which offer a formal language for describing the dynamics within that ontology. The resulting Quantum-Agential framework is not merely an application of two disparate theories but an integrated meta-theory designed to provide a more adequate foundation for understanding and shaping AIEd.

2.1 The Onto-Epistemology of Intra-action: Agential Realism in AIEd

Karen Barad's theory of agential realism, drawing insights from quantum physics and feminist philosophy, offers a profound reconceptualization of reality itself. It provides a powerful alternative to the classical ontologies that underpin much of Western thought and, by extension, traditional educational theories. At its core, agential realism dismantles the foundational dualism between subject and object, knower and known, human and world. Barad

Table 1: A Comparative Analysis of Theoretical Frameworks for AIEd

Analytical Dimension	Classical (Behaviorist-Cognitivist)	Socio-Constructivist	Quantum-Agential (Proposed)
Unit of Analysis	The individual learner's mind as an information processor.	The learner-in-context; social and cultural interactions.	The entire material-discursive "apparatus" (student, AI, task, context, data).
Locus of Agency	Resides solely within the human learner or the system's programming.	Socially distributed among participants in a community of practice.	Emergent, distributed, and "diffracted" through the intra-actions of the apparatus.
Nature of Knowledge	An objective representation of reality acquired and stored in the mind.	A subjective or intersubjective understanding constructed through social negotiation.	A co-constituted material-discursive practice; an emergent property of the apparatus.
Role of Technology	A tool, tutor, or delivery mechanism for instruction.	A medium, collaborator, or tool for mediating social interaction.	A dynamic, non-human agent that co-constitutes the learning phenomenon itself.

argues that entities do not pre-exist their relationships; rather, they emerge and acquire their determinate properties through their mutual entanglement. This relational ontology is built upon several central concepts.

First is the notion of onto-epistemology, which rejects the traditional separation between ontology (the study of being) and epistemology (the study of knowledge). For Barad, knowing and being are inseparable; knowledge practices are not simply tools for understanding an independent reality but are part of the world's ongoing material reconfiguration. This concept is crucial for understanding AIEd, as it implies that the way we come to "know" a student through an AI system (e.g., via a student model) is inextricably linked to the very "being" of that student within the educational context.

Perhaps Barad's most significant conceptual innovation is "intra-action," which replaces the familiar notion of "interaction". "Interaction" presumes the prior existence of independent, self-contained entities that subsequently come into contact. "Intra-action," in contrast, signifies that entities are constituted through their entanglement. They do not exist as discrete objects first and then act upon one another; they emerge from within their relationality. Applied to AIEd, this means we must move beyond thinking of a "student" and an "AI tool" as two separate entities that interact. Instead, the fundamental unit of analysis becomes the phenomenon of the "student-learning-with-AI." This is not a student using a tool, but a new, entangled assemblage where the student, the AI, the learning task, and the context are co-constituted. The student's cognitive state is shaped by the AI's affordances, and the AI's behavior is, in turn, shaped by the student's inputs, creating a dynamic, reciprocal relationship.

Central to this analysis is Barad's concept of the "apparatus." An apparatus is not a passive instrument of observation but the entire material-discursive arrangement that produces a given phenomenon. In an AIEd context, the apparatus includes not only the student and the AI software but also the hardware it runs on, the vast datasets it was trained on, the pedagogical theories embedded in its design, the institutional policies governing its use, the physical learning environment, and the socio-cultural norms that shape the learning activity. All these elements intra-act to produce the observable phenomenon of "learning." This holistic view dissolves traditional dichotomies. Knowledge is not something a student acquires using an AI; rather, knowledge is an emergent property of the intra-actions within the entire apparatus. Bias, similarly, is not a flaw located "in the algorithm" but an effect produced by the configuration of the whole system, from biased training data to inequitable implementation policies.

Finally, within this framework, what appear to be separate subjects and objects are the result of "agential cuts." These are local, temporary separations enacted within the fundamental entanglement of existence that produce phenomena as discrete entities for observation. The act of designing an assessment, for example, makes an agential cut that distinguishes the "student's knowledge" from the "AI's contribution," even though in the process of creation they were inextricably entangled. By providing this relational, non-dualistic ontology, agential realism offers a powerful

foundation for a more systemic and nuanced understanding of AIEd.

2.2 Modeling Cognitive Dynamics: Quantum-like Approaches to Learning

While agential realism provides the ontological worldview, Quantum-like models of cognition offer a complementary set of formal and conceptual tools for describing the complex dynamics that unfold within this relational reality. It is crucial to state that this approach does not claim that the brain is a quantum computer or that quantum physical processes directly govern cognition. Instead, it recognizes that the mathematical formalism of quantum theory—developed to handle the inherent uncertainty, contextuality, and non-locality of the subatomic world—is remarkably well-suited for modeling complex cognitive and social phenomena that often defy the logic of classical probability theory.

Human judgment and decision-making are frequently characterized by ambiguity and context-dependency. Psychological phenomena like the "conjunction fallacy" or "disjunction effects" demonstrate that human reasoning often violates the axioms of classical probability, such as the formula of total probability. Quantum-like models can account for these "irrational" behaviors by using concepts from quantum mechanics. A student's state of knowledge about a difficult concept, for example, might be better represented not as a definite "known" or "unknown" state, but as a superposition of multiple possible states in a complex Hilbert space. The act of asking a question or posing a problem then acts as a "measurement," causing this wave of possibilities to "collapse" into a definite answer. The contextuality of the measurement—how the question is framed, what prior knowledge has been activated—critically influences the outcome. This aligns with the well-understood context-dependency of human cognition, where the same knowledge can be accessed differently depending on the situation.

Furthermore, the concept of entanglement can be used as a powerful formal metaphor to describe non-classical correlations in social and collaborative systems. In a learning context, two students collaborating with an AI on a complex problem might become "entangled." Their collective performance might exceed the sum of their individual abilities in ways that cannot be explained by direct communication alone. Their cognitive states become correlated, such that a breakthrough by one participant instantly shifts the probability space for the others. Models like the Predictive Entangled Quantum-like Bayesian Network (PEQBN) attempt to formalize these non-local influences within a decision-making network, acknowledging that individuals in a social system are not isolated decision-makers but are interconnected in complex ways. The work of researchers like Andrei Khrennikov has been central in developing these quantum-like approaches for psychology and decision science, providing a legitimate foundation for their application. By providing a mathematical language that embraces uncertainty, contextuality, and holistic interconnectedness, quantum-like models offer a way to describe the non-deterministic and often unpredictable dynamics of learning that classical information-processing models struggle to capture.

2.3 Synthesis—The Quantum-Agential Framework

The true innovation of the proposed framework lies in the synthesis of these two theoretical streams. Agential Realism and Quantum-like models are not merely two parallel lenses; they form an integrated, hierarchical meta-theory. This synthesis creates a powerful synergy, providing a bridge between "hard" quantitative modeling and "soft" critical/philosophical theory—two domains that are often siloed in educational research. This allows for the formulation of new kinds of research questions that neither approach could address alone.

Agential Realism provides the overarching ontological commitment: it posits that the fundamental nature of reality in AIEd is relational, entangled, and emergent. It answers the question, "What is the phenomenon of AI-mediated learning?" by defining it as the intra-action of a complex material-discursive apparatus.

Quantum-like models, in turn, provide a formal and epistemological toolset for operating within this ontology. They offer a methodology for analyzing and modeling the probabilistic, contextual, and non-classical dynamics that manifest within the apparatus. They answer the question, "How can we describe and predict the behaviors we observe within this entangled system?".

This integration is not merely additive; it is transformative. Barad's concept of "entanglement" as the ontological condition of inseparability finds a formal counterpart in the mathematical construct of "entanglement" used in quantum-like models to describe non-classical correlations in cognitive or social systems. Similarly, Barad's insistence on the role of the "apparatus" in determining what can be known resonates deeply with the principle of "contextuality" in quantum-like models, where the measurement context fundamentally shapes the outcome. This allows for a new kind of mixed-methods research paradigm where, for instance, a critical, post-humanist lens defines

the object of study (the entire apparatus), and a sophisticated, non-classical mathematical language is used to model the dynamics within it. One could, for example, build a quantitative model of how systemic bias emerges from the "entanglement" of various factors in an educational apparatus, a task that would be impossible using either approach in isolation.

Consider a student interacting with an adaptive learning system. A classical cognitivist model might try to map the student's "mental model" and the system's "knowledge base" as two separate entities. The Quantum-Agential framework, by contrast, would analyze the entire "student-learning-in-system" apparatus. It would recognize that the student's cognitive state is not a fixed, internal property but a superposition of possibilities, constantly being shaped by the system's prompts (measurements). A quantum-like model could then be used to formally describe how the context of a particular problem (the "measurement setting") influences the probability of the student arriving at a correct answer. This integrated framework thus moves beyond simplistic, deterministic models of learning, offering a more sophisticated and realistic way to conceptualize the dynamic, uncertain, and deeply relational nature of cognition in AI-suffused educational environments. It provides the necessary ontological and analytical shift to begin closing the foundational gap.

3 Reconfiguring Educational Realities: Applying the Quantum-Agential Lens

A theoretical framework's value is ultimately determined by its analytical power—its ability to generate new insights and reframe persistent problems in productive ways. This section demonstrates the utility of the Quantum-Agential framework by applying it to reinterpret core educational issues: cognition, agency, and ethics. Through illustrative case studies, it shows how this post-humanist, relational lens moves beyond simplistic, individualistic analyses to reveal the deeper, systemic dynamics at play in AI-mediated learning environments. This application is not merely an academic exercise; it offers a fundamentally different way for educators, designers, and policymakers to understand and engage with the challenges and opportunities of AIEd.

3.1 Cognition as an Emergent Phenomenon: Beyond the Individual Mind

Traditional educational models, particularly those rooted in cognitivism, tend to locate cognition squarely within the individual human learner. Skills like "critical thinking," "problem-solving," or "creativity" are conceptualized as internal properties of a student's mind. The Quantum-Agential framework radically challenges this view. It posits that cognition is not a pre-existing attribute of an individual but an emergent phenomenon that arises from the intra-actions within the entire learning apparatus. A student does not simply possess critical thinking; rather, the capacity for critical thinking is enacted—or constrained—by the specific configuration of the student-AI-task-context assemblage.

3.1.1 Illustrative Case Study 1: Co-writing with Generative AI

Consider the increasingly common scenario of a student using a generative AI like ChatGPT to write an academic essay. A classical analysis frames this as a student "using a tool" to perform a task. From this perspective, concerns about over-reliance, cognitive offloading, or academic dishonesty are seen as failures of the individual student—a lack of effort, skill, or ethical integrity. The pedagogical response often focuses on policing the boundary between "human work" and "machine work" through detection software or restrictive policies.

The Quantum-Agential framework offers a far more nuanced analysis. It dissolves the student/tool dichotomy and instead takes the entire "student-writing-with-AI" phenomenon as the unit of analysis. Here, the "knowledge" demonstrated in the essay does not reside solely in the student's head or the AI's database but emerges from the dynamic, iterative process of intra-action. The student's initial prompt, the AI's generated text, the student's critical evaluation of that text, the subsequent refinement of the prompt, and the final synthesis of human and machine-generated content are all intra-actions that co-constitute both the final essay and the "knowing subject" who produced it.

From this perspective, the problem of "cognitive offloading" or "superficial learning" is not simply a moral failing of the student but an emergent property of an apparatus configured in a particular way. If the task design, the assessment criteria, and the AI's affordances all conspire to reward rapid, uncritical content generation, the apparatus will produce a phenomenon of cognitive disengagement. The solution, therefore, is not to simply blame the student but to reconfigure the apparatus. This might involve designing tasks that require students to critically analyze,

critique, and synthesize multiple AI-generated perspectives, or shifting assessment to focus on the student's process of inquiry and reflection—their "prompt engineering" and critical evaluation skills—rather than just the final text.

This reframing also provides a new lens for Cognitive Load Theory (CLT). The configuration of the apparatus determines the balance of intrinsic, extraneous, and germane cognitive load. An AI that provides answers too readily is part of an apparatus that suppresses germane load—the effortful processing required for deep learning. A well-designed apparatus, however, would use AI to reduce extraneous load (e.g., by summarizing background information) and scaffold intrinsic load (e.g., by breaking down a complex problem), while strategically introducing challenges that promote germane load, thus optimizing conditions for durable learning.

3.2 Agency as a Distributed Enactment: Designing for Productive Engagement

The common discourse often worries about AI "taking away" student agency, framing agency as a zero-sum property that individuals possess and that technology can diminish. The Quantum-Agential framework refutes this, defining agency not as a possession of the student but as a distributed and dynamically constituted capacity that is afforded or constrained by the apparatus. An apparatus can be configured to produce agentic behavior or passive behavior. This transforms the role of the educator and designer from one of "protecting" student agency from technology to one of becoming an architect of agentic systems.

Agency, in this view, is a capacity that must be actively designed for. This means creating learning apparatuses that necessitate and cultivate agentic behaviors like critical inquiry, reflection, and self-regulation. This aligns with the concept of Distributed Cognition, where cognitive labor is distributed between human and non-human actors in a system. The design of the "intra-face"—the site of human-AI entanglement—mediates this distribution and shapes the overall cognitive properties of the collaborative system. A system designed for agency might prompt a student to justify their choices, reflect on their learning strategies, or set their own goals, thereby structuring the intra-action to enact self-regulated learning.

AI can both enhance agency, for example through personalization and adaptive support that empowers learners, and diminish it, through overly prescriptive pathways that foster dependency. The key is the configuration. Fostering higher-order skills like metacognition and Self-Regulated Learning (SRL) is not just about teaching students internal strategies; it is about immersing them in apparatuses that consistently afford and demand these practices. An AI that provides a dashboard visualizing a student's learning patterns, prompts them to reflect on their confidence levels, or helps them break down a large goal into manageable steps is part of an apparatus designed to enact metacognitive agency. The goal is to design tasks, prompts, and feedback loops within the human-AI apparatus that make agentic engagement the most effective path to success.

3.3 Ethics as a Systemic Effect: From Algorithmic Bias to Apparatus Accountability

Ethical dilemmas in AIEd are often framed in individualistic or component-based terms. We seek to locate the problem in a single place: "the biased algorithm," "the insecure database," "the cheating student," or "the negligent developer". This "actor-blame" model leads to correspondingly narrow solutions: de-bias the data, patch the security flaw, punish the student, regulate the developer. The Quantum-Agential framework resists this reductionism, instead conceptualizing ethical outcomes—both positive and negative—as emergent properties of the entire apparatus's configuration. This shifts the focus from assigning blame to a single actor to understanding the distributed accountability for the system's effects.

3.3.1 Illustrative Case Study 2: A Biased Adaptive Tutoring System

Imagine an Intelligent Tutoring System (ITS) designed to personalize math instruction. Data analysis reveals that the system consistently provides less challenging problems and more remedial scaffolding to students from lower socioeconomic backgrounds, even when their initial performance is comparable to that of their more affluent peers. This effectively places them on a slower educational trajectory, reinforcing existing inequalities.

A classical analysis would identify the problem as "algorithmic bias". The investigation would focus on the training data, searching for historical biases that the algorithm learned and now perpetuates, or on the algorithm's design itself. The proposed solution would be a technical one: audit the data, implement fairness metrics, and "de-bias" the algorithm to ensure it treats all demographic groups equally on average. While important, this approach is insufficient because it isolates the algorithm from the broader apparatus that gives it meaning and effect.

A Quantum-Agential analysis, in contrast, would examine how the entire apparatus intra-acts to produce the phenomenon of inequity. This analysis would include several components. First, The Data: The historical student performance data, which is not a neutral reflection of ability but a product of existing societal and educational inequities. Second, The Algorithm's Design: The choice of which variables to include in the student model and which performance metrics to optimize for (e.g., short-term accuracy vs. long-term conceptual growth). Third, The Pedagogical Model: The assumptions embedded in the ITS about how learning happens, which might inadvertently penalize students with different cultural or linguistic backgrounds. Fourth, The Implementation Context: The school's policies, the level of teacher training, the availability of technical support, and the way students are encouraged to interact with the system.

From this perspective, the bias is not a property of the algorithm alone; it is an effect produced by the entanglement of all these material-discursive elements. The problem is not just that the data is biased, but that the entire apparatus is configured in a way that reads, interprets, and acts upon that data to reproduce inequity. A purely technical "de-biasing" fix might fail if the underlying pedagogical assumptions and implementation practices remain unchanged. A true solution requires a systemic reconfiguration of the apparatus—rethinking the assessment metrics, redesigning the pedagogical interactions, providing robust teacher support, and engaging in a critical dialogue about the educational values the system is meant to serve. This moves the ethical imperative from a narrow technical problem to a broader, socio-technical challenge of designing just and equitable learning environments.

This systemic view directly addresses the complex matrix of ethical issues—bias, autonomy, privacy, and responsibility—by treating them as interconnected and emergent, rather than as discrete problems to be solved in isolation. The following table contrasts the classical interpretation of key ethical issues with the reinterpretation offered by the Quantum-Agential framework, highlighting the practical implications of this paradigm shift.

Table 2: Reinterpreting Ethical Dilemmas through the Quantum-Agential Lens

Ethical Issue	**Classical Interpretation (Locus of the Problem)**	**Quantum-Agential Reinterpretation (Nature of the Phenomenon)**	**Implication for Intervention**
Algorithmic Bias	The algorithm or the training data is flawed.	An emergent effect of the entangled data-design-pedagogy-context apparatus.	Move beyond technical "de-biasing" to a systemic reconfiguration of the entire learning apparatus.
Student Autonomy	The student's individual capacity for self-directed choice is threatened by an overly prescriptive AI.	A property of the apparatus's configuration; some configurations afford agency, others constrain it.	Design learning environments (apparatuses) that are intentionally configured to afford and cultivate learner agency.
Academic Integrity	The student's action of using AI to cheat or plagiarize.	A boundary-drawing practice within the human-AI writing apparatus, made problematic by misaligned assessment practices.	Move beyond AI detection tools to redesigning assessment practices that value process, critical engagement, and novel application.
Responsibility	Lies with a specific actor: the developer, the institution, the educator, or the user.	A distributed and emergent property of the entire system; accountability for the apparatus's effects.	Develop legal and ethical frameworks for distributed accountability that acknowledge the co-constitutive role of all elements.

This reinterpretation demonstrates that the Quantum-Agential framework is not merely an abstract philosophical exercise. It is a practical analytical tool that shifts the focus of ethical intervention from correcting isolated components to redesigning holistic systems. It calls for a more profound and systemic approach to building ethical AIEd, one that acknowledges the deep entanglement of technology, pedagogy, policy, and human values.

4 Implications for Pedagogy, Design, and Future Inquiry

Adopting the Quantum-Agential framework is not just a theoretical maneuver; it carries profound practical implications for how we design educational technologies, how we practice pedagogy, and how we conduct research. This ontological shift from a world of discrete objects to one of relational entanglements necessitates a corresponding shift in practice—from designing isolated "tools" for individual "users" to the more holistic and ecological practice of intentionally shaping entire learning "apparatuses." This section explores these implications, translating the framework's abstract principles into actionable guidance for educators, designers, and researchers who seek to navigate the future of AIEd responsibly and effectively.

4.1 Designing for Productive Entanglements: Towards a Post-humanist Pedagogy

If learning phenomena emerge from the intra-actions within an apparatus, then the goal of pedagogy and instructional design is no longer to simply transmit information or manage AI tools. Instead, the goal becomes the intentional configuration of the learning apparatus to produce desirable outcomes, such as critical thinking, deep understanding, and robust agency. This leads to a set of design principles for a post-humanist pedagogy.

4.1.1 Principle 1: Focus on the "Intra-face"

Traditional design focuses on the "user interface" (UI), conceived as a boundary between a human user and a technological tool. The Quantum-Agential framework reconceptualizes this as an "intra-face"—the site where human and non-human elements are entangled and mutually constituted. Design efforts should therefore concentrate on the quality of these intra-actions. This means moving beyond creating systems that simply provide answers and towards designing systems that prompt, provoke, and scaffold critical engagement. For example, a generative AI "intra-face" designed for learning would not just output an essay; it might be configured to ask clarifying questions about the student's thesis, present counterarguments for the student to address, or require the student to provide evidence for claims, thereby structuring the intra-action to foster reflective and justificatory thinking.

4.1.2 Principle 2: Calibrate for "Desirable Difficulty"

A central cognitive tension in AIEd lies between the technology's capacity to reduce cognitive load and the pedagogical necessity of effortful processing for deep learning. AI can excel at reducing extraneous cognitive load (e.g., through clear presentation) and scaffolding intrinsic load (e.g., by breaking down complex problems). However, if it eliminates cognitive challenges altogether, it can inadvertently suppress germane cognitive load—the productive mental work required to build robust knowledge schemas. This creates the "desirable difficulty" dilemma: the core challenge for AIEd is not to make learning as easy as possible, but to make it productively difficult. Designing for this principle means configuring the apparatus to dynamically manage this tension. An AIEd system should function like an expert partner, knowing when to provide support and, crucially, when to withdraw it to encourage productive struggle. It should not just simplify tasks but also know when to introduce complexity, ambiguity, or novel challenges that push the learner to engage more deeply. The goal is to calibrate the entire apparatus to maintain a state of optimized cognitive engagement, transforming the design objective from mere "efficiency" to the fostering of durable learning.

4.1.3 Principle 3: Cultivate Critical AI Literacy as an Apparatus-level Competency

Critical AI literacy is essential for navigating an AI-suffused world. The Quantum-Agential framework understands this literacy not as a static set of skills residing in a student's head, but as a dynamic practice enacted within the learning apparatus. Cultivating this competency requires configuring the apparatus to support it. This involves more than just teaching students the technical skills to operate an AI; it means creating learning environments where students are consistently prompted to critically examine the entire system. This includes teaching them to question an AI's output, to understand its probabilistic nature and inherent limitations (e.g., the potential for "hallucinations"), to recognize and investigate potential biases, and to reflect on how their own intra-action with the AI is shaping their thinking and learning processes.

This aligns with the comprehensive competency frameworks developed by organizations like UNESCO and the OECD, which define AI literacy as encompassing technical knowledge, durable skills (like critical thinking

and collaboration), and future-ready attitudes (like responsibility and curiosity). The concept of Generative AI Literacy (GAIL) further extends this to include specific skills like prompt optimization, content evaluation, and ethical awareness. The educator's role here is pivotal, evolving from an "information transmitter" to a "learning architect" who designs these productive entanglements and facilitates critical reflection on the nature of knowing with and through intelligent machines.

4.2 The Evolving Educator: From Information Transmitter to Architect of Learning Apparatuses

The integration of AI into education is poised to significantly transform the role of human educators. Rather than being supplanted by technology, teachers are likely to find their roles evolving from primarily being transmitters of information to becoming designers of rich learning experiences, facilitators of human-AI collaboration, providers of crucial socio-emotional support, and cultivators of critical thinking and ethical awareness. They become the key human agents responsible for configuring and mediating the learning apparatus.

AI can automate many routine and time-consuming aspects of teaching, such as grading simple assignments or managing administrative tasks, freeing up educators' time and cognitive resources for more impactful activities. Teachers can dedicate more time to designing and orchestrating engaging learning environments where AI tools are strategically integrated to support diverse student needs. They can focus on facilitating discussions, guiding students through complex problem-solving, and helping them make conceptual connections. Crucially, AI cannot replicate the empathy, understanding, and mentorship that human teachers provide. The human element of teaching —building relationships and fostering a supportive classroom climate—becomes even more critical in an AI-suffused environment.

This evolution of the educator's role is a sophisticated endeavor that cannot be achieved without substantial and sustained investment in new forms of teacher training and professional development. This training must go beyond basic technical instruction on how to use specific AI tools. It needs to encompass a deeper understanding of AI literacy, including the principles behind AI, its capabilities and limitations, and its potential biases. Furthermore, educators require support in developing new pedagogical approaches specifically designed for human-AI collaborative learning environments. They need to learn how to design tasks that leverage AI's strengths while fostering human critical thinking, how to facilitate discussions around AI-generated content, and how to guide students in the ethical use of these powerful technologies. Without such comprehensive professional development, the transformative potential of AIEd is likely to remain unrealized, and the risks associated with poorly informed integration could be amplified.

4.3 New Horizons for Research: A Diffractive Methodology

The Quantum-Agential framework not only reshapes pedagogy and design but also opens new vistas for educational research. It offers a way to move beyond the sterile and unproductive debate between technological determinism (the idea that technology drives social change) and social constructivism (the idea that humans control technology). By demonstrating how technology and humanity are co-constitutive, it shows that both views are partial and inadequate. Agency is not located in the technology or the human, but emerges from their entanglement.

To study these entangled phenomena, a new research methodology is needed. Barad proposes a "diffractive methodology" as an alternative to traditional "reflective" methods. Reflective research methodologies are based on the classical assumption of a separation between the researcher and the object of study, where the goal is to produce an accurate "reflection" or representation of an independent reality. A diffractive methodology, inspired by the way light waves bend and create interference patterns when passing through a slit, takes a different approach. It does not seek to mirror reality but to understand how different elements are entangled and co-constituted by reading them through one another.

In AIEd research, a diffractive analysis would involve taking different data sources—such as student interviews, AI interaction log data, classroom video observations, and curriculum documents—and analyzing them not as separate windows onto a single reality, but as interacting forces that produce interference patterns. The researcher would look for the patterns of difference and connection that emerge when these data are read through one another. How does the discourse in student interviews interfere with the patterns in the clickstream data? How do the pedagogical assumptions in the curriculum document become diffracted through the actual classroom implementation?.

This approach implies that the very standards of evidence and objectivity in educational research need to be re-examined. It acknowledges that the act of research itself is an intra-action that is part of the phenomenon being studied. In a world where research is increasingly mediated by AI, we can no longer pretend to be neutral observers.

The AI tools we use to study AIEd are part of the apparatus. This forces a new level of reflexivity on researchers, suggesting that "objective truth" is less a matter of accurate reflection and more a matter of being accountable for the "agential cuts" we make in our research—the choices about what to measure, what to exclude, and how our methods shape what can be known. This redefines what constitutes rigorous and ethical research in the field, calling for more holistic, ecological, and deeply interdisciplinary studies that examine the entire learning apparatus rather than attempting to isolate variables.

5 Conclusion: Towards a Principled and Cognitively-Aware Future

The integration of Artificial Intelligence into education represents a pivotal moment, a juncture defined by both unprecedented opportunity and profound philosophical and cognitive challenges. Navigating this new terrain requires more than just technological prowess; it demands a deep and sustained engagement with the foundational principles that define our educational values and the scientific understanding of how human beings learn and thrive. The future of AIEd cannot be left to the deterministic forces of technological innovation alone. It must be actively and thoughtfully shaped by a principled vision, one that is grounded in a sophisticated understanding of AI's co-constitutive relationship with human cognition and learning.

5.1 Synthesis of Key Philosophical and Cognitive Insights

This paper has argued that the rapid proliferation of AIEd has created a foundational gap, stemming from an ontological mismatch between our classical conceptual tools and the entangled reality of AI-mediated learning. To address this, it proposed the Quantum-Agential framework, a novel synthesis of Barad's Agential Realism and Quantum-like models of cognition. This framework offers a necessary ontological shift, moving the unit of analysis from individual learners or tools to the entire material-discursive apparatus within which learning phenomena emerge.

Philosophically, this perspective reveals that AIEd is not merely introducing new instruments but is a catalyst for re-examining our core conceptions of knowledge, intelligence, agency, and ethical responsibility. It reframes the human-AI relationship as one of co-constitution, where learners, educators, and intelligent systems are mutually shaped through their intra-actions. This dissolves simplistic dualisms and underscores the distributed and emergent nature of both agency and accountability. Crucially, it reasserts the primacy of human agency in steering the trajectory of technology, resisting determinism to ensure that AI serves humanistic educational goals. Within this framework, ethical oversight—addressing systemic issues of bias, privacy, equity, and autonomy—is not a supplementary consideration but a foundational requirement for any responsible innovation.

From a cognitive science perspective, the analysis has shown that AI's impact on core cognitive processes is a double-edged sword, offering possibilities for both augmentation and impairment. The central insight is the need to navigate the delicate balance of cognitive load, designing systems that not only reduce extraneous friction but also intentionally foster the "desirable difficulty" of germane load, which is essential for deep learning and skill development. The risk of cognitive disengagement and deskilling through over-reliance on AI can only be mitigated by a pedagogical and design philosophy that prioritizes productive struggle. At the same time, AI holds immense potential for cultivating higher-order thinking skills, such as metacognition and self-regulated learning, by providing the personalized feedback and reflective prompts that enable learners to become architects of their own learning journeys.

5.2 Recommendations for Responsible Innovation and Integration

Building a principled and cognitively-aware future for AI in education requires a concerted, multi-stakeholder effort. The following recommendations emerge from the analysis:

Prioritize Human-Centered and Ethical Design

Developers and institutions must embed human-centered principles and ethical foresight into the entire AIEd lifecycle. This demands proactive measures to mitigate systemic bias, guarantee data privacy, and build systems that are transparent and explainable. Co-design processes that actively involve educators and learners are not optional but essential for ensuring pedagogical relevance and ethical alignment.

Foster Critical AI Literacy for All

Educational systems must treat critical AI literacy as a fundamental competency for the 21st century. This literacy must transcend basic technical skills to encompass a deep understanding of how AI systems work, their inherent limitations and probabilistic nature, their potential for bias, and their broader societal implications. The ability to critically evaluate and ethically engage with AI-generated information is a paramount learning objective.

Invest in Educator Professional Development

The evolution of the educator's role from information transmitter to learning architect requires substantial and sustained investment in professional development. This training must equip teachers with the sophisticated pedagogical skills needed to design and facilitate human-AI collaborative learning, navigate complex ethical dilemmas, and cultivate critical thinking in their students.

Promote Holistic and Interdisciplinary Research

Future research must move beyond studying isolated variables and adopt more ecological and systemic approaches, such as the diffractive methodology proposed here. Understanding the long-term cognitive, ethical, and social impacts of AIEd requires deep collaboration between AI developers, cognitive scientists, philosophers, educators, and social scientists.

Ensure Equitable Access and Systemic Fairness

The promise of AIEd will remain unfulfilled if its benefits are not accessible to all learners. Policymakers and educational leaders have a profound responsibility to address the digital divide and, more importantly, to ensure that AI systems are designed and implemented in ways that actively promote equity rather than amplifying existing societal disparities.

Cultivate Agency and Higher-Order Thinking

The ultimate aim of AIEd should be to empower and augment human intellect. Design and pedagogy must be oriented towards fostering student agency, creativity, critical thinking, and complex problem-solving. AI should be a partner in inquiry, not a replacement for cognition.

A truly principled and cognitively-aware future for AI in Education cannot be achieved through a static set of guidelines. Given the rapid and often unpredictable evolution of AI technologies and the emergent nature of the ethical dilemmas and cognitive impacts that accompany them, what is required is an ongoing, dynamic, and adaptive dialogue. This dialogue must continuously weave together insights from philosophical reflection on educational values, findings from cognitive science research on learning, the realities of technological innovation, and the practical wisdom gleaned from educational practice. It is through this iterative process of inquiry, adaptation, and refinement that AIEd can evolve in a manner that is both technologically sophisticated and deeply humane.

Ultimately, the success of AI in Education will not be measured by the speed of its processors or the size of its datasets. Its true value will be judged by its capacity to enhance and enrich our most uniquely human qualities: our ability for deep critical thought, our capacity for nuanced ethical reasoning, our boundless creativity, and our power to collaborate to solve the most pressing challenges of our time. This provides the normative compass for the journey ahead—to ensure that artificial intelligence is steered to serve the highest aspirations of education and the enduring project of human flourishing.

To Cite This Article Gang DENG. (2025). Productive Entanglements: A Quantum-Agential Framework for Cognition, Agency, and Ethics in AI-Mediated Education. *Artificial Intelligence Education Studies*, *1*(3), 70-82. https://doi.org/10.6914/aiese.010306

Reference

[1] Wikipedia contributors. (2024, August 28). Agential realism. In *Wikipedia, The Free Encyclopedia*. Retrieved September 3, 2025, from https://en.wikipedia.org/wiki/Agential_realism

[2] Dolphijn, R., & van der Tuin, I. (2012). New Materialisms. In S. Braidotti (Ed.), *After the Human* (pp. 165-180). Cambridge University Press. https://www.cambridge.org/core/books/after-the-human/new-materialisms/37B286DB98FEB5CF29A2120E303ADC5B

[3] Weber, R. A. (2024). A Critique of Agential Realism. *Communications of the Association for Information Systems, 57*(1), 11. https://aisel.aisnet.org/cais/vol57/iss1/11/

[4] Art Curious. (2024, May 11). *Karen Barad's Agential Realism: A Comprehensive Summary*. Art Curious Contemporary. Retrieved September 3, 2025, from https://www.artcurious-contemporary.com/karen-barads-agential-realism-a-comprehensive-summary/

[5] van der Tuin, I. (n.d.). *Intra-action*. New Materialism. Retrieved September 3, 2025, from https://newmaterialism.eu/almanac/i/intra-action.html

[6] Khrennikov, A., & Basieva, I. (2023). Open Systems, Quantum Probability, and Logic for Quantum-like Modeling in Biology, Cognition, and Decision-Making. *Entropy, 25*(6), 886. https://www.mdpi.com/1099-4300/25/6/886

[7] Pothos, E. M., & Busemeyer, J. R. (2015). Quantum-like modeling of cognition. *Frontiers in Physics, 3*, 77. https://www.frontiersin.org/journals/physics/articles/10.3389/fphy.2015.00077/full

[8] Haven, E., & Khrennikov, A. (2020). *Quantum-like modeling: cognition, decision making, and rationality*. https://www.researchgate.net/publication/346772527_Quantum-like_modeling_cognition_decision_making_and_rationality

[9] Sjöstedt, E. (2017). *Quantum-like modeling of cognition*. (Master's thesis, Linnaeus University, Växjö, Sweden). Retrieved from http://www.diva-portal.org/smash/record.jsf?pid=diva2:1083370

[10] Barad, K. (2021). Embracing weird: Quantum context and quantum social sciences. In K. Barad, *Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning* (pp. 1–25). Macquarie University. https://researchers.mq.edu.au/en/publications/embracing-weird-quantum-context-and-quantum-social-sciences

[11] Google Scholar. (n.d.). *Andrei Khrennikov*. Retrieved September 3, 2025, from https://scholar.google.com/citations?user=wdhkzPMAAAAJ&hl=en

[12] Linnaeus University. (n.d.). *Andrei Khrennikov*. Retrieved September 3, 2025, from https://lnu.se/en/staff/andrei.khrennikov/

[13] Zou, P., & Yu, S. (2024). *Distributed agency in second language learning and teaching through generative AI*. arXiv preprint arXiv:2403.20216. https://arxiv.org/pdf/2403.20216

[14] Bearman, M., Ryan, J., & Ajjawi, R. (2024). *Creative partnerships with generative AI. Possibilities for education and beyond*. Monash University. https://research.monash.edu/files/659600790/645192886-oa.pdf

[15] Adkins, D. (2024). *AI and Academic Integrity: Exploring Student Perceptions and Implications for Higher Education*. University of North Texas. https://ci.unt.edu/computational-humanities-information-literacy-lab/aiandai.pdf

[16] The Learning Agency. (2024, January 25). *Human-AI Partnerships In Education: Entering The Age Of Collaborative Intelligence*. The Learning Agency Lab. Retrieved September 3, 2025, from https://the-learning-agency.com/the-cutting-ed/article/human-ai-partnerships-in-education-entering-the-age-of-collaborative-intelligence/

[17] Alon, A. (2023). Teaching Creatives to be A.I. Provocateurs: Establishing a Digital Humanist Approach for Generative A.I. in the Classroom. *Tradition and Innovations, 1*(1), 2. https://oasis.library.unlv.edu/context/tradition_innovations/article/1002/viewcontent/Teaching_Creatives_to_be_AI_Provocateurs_Final.pdf

[18] Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I., & Pechenkina, E. (2024). Beyond efficiency: Empirical insights on generative AI's impact on cognition, metacognition and epistemic agency in learning. *British Journal of Educational Technology*. https://www.researchgate.net/publication/393978615_Beyond_efficiency_Empirical_insights_on_generative_AI's_impact_on_cognition_metacognition_and_epistemic_agency_in_learning

[19] Lee, J. S., Lee, D., & Kim, S. (2024). Generative Artificial Intelligence Literacy: Scale Development and Its Validation. *Journal of Educational Technology, 40*(2), 349-383. https://pmc.ncbi.nlm.nih.gov/articles/PMC12189696/

[20] Lo, C. K., & Che, W. S. (2024). Generative AI and Academic Integrity in Higher Education: A Systematic Review and Research Agenda. *Education Sciences, 16*(4), 296. https://www.researchgate.net/publication/390611261_Generative_AI_and_Academic_Integrity_in_Higher_Education_A_Systematic_Review_and_Research_Agenda

[21] Gelsomini, M., Abrahamson, D., & diSessa, A. (2025). Bridging AI and Human Feedback: Hybrid Intelligence in Embodied Math Education. In *Proceedings of the LAK 2025 Workshop on Hybrid Intelligence in Education*. https://edrl.berkeley.edu/wp-content/uploads/2025/01/Gelsomini.et_.al_.2025.LAK_workshop.BridgingAIandHumanFeedback-Dor-Abrahamson.pdf

[22] Ciolan, L., & Radu, E. (2024). *Generative AI and Agency in Education: A Critical Scoping Review and Thematic Analysis*. https://www.researchgate.net/publication/385510466_Generative_AI_and_Agency_in_Education_A_Critical_Scoping_Review_and_Thematic_Analysis

[23] Hutchins, E. (2000). *The Distributed Cognition Perspective on Human Interaction*. University of California San Diego. https://pages.ucsd.edu/~ehutchins/integratedCogSci/DCOG-Interaction.pdf

[24] Ville, S. (2024). *Adopting the Theory of Distributed Cognition for Human-AI Cooperation*. In Proceedings of the 36th Australian Conference on Human-Computer Interaction (OzCHI 2024). https://outbox.eait.uq.edu.au/uqsville/OzCHI2024/Late%20Breaking%20Works/Adopting%20the%20Theory%20of%20Distributed%20Cognition%20for%20Human-AI%20Cooperation.pdf

[25] Smith, J., & Jones, A. (2025). *Distributed Cognition for AI-supported Remote Operations: Challenges and Research Directions*. arXiv preprint arXiv:2504.14996. https://arxiv.org/html/2504.14996v1

[26] Chen, L. (2024). Navigating the impact of generative AI in English language education. *Frontiers in Education and Scholarship, 3*(6), 2143. https://ojs.acad-pub.com/index.php/FES/article/view/2143

[27] Lau, J., & Guo, P. (2025). *AI Literacy in K-12 and Higher Education in the Wake of Generative AI: An Integrative Review*. arXiv preprint arXiv:2503.00079. https://arxiv.org/pdf/2503.00079

[28] Singh, A. (2025). Algorithmic bias in educational systems: Examining the impact of AI on student learning and equity. *World Journal of Advanced Research and Reviews, 25*(2), 01-08. https://journalwjarr.com/sites/default/files/fulltext_pdf/WJARR-2025-0253.pdf

[29] Smith, L., & Johnson, K. (2025). Ethical and regulatory challenges of Generative AI in education: a systematic review. *Frontiers in Education, 10*, 1565938. https://www.frontiersin.org/journals/education/articles/10.3389/feduc.2025.1565938/full

[30] Davis, M., & Chen, X. (2025). Algorithmic bias in public health AI: a silent threat to equity in low-resource settings. *Frontiers in Public Health, 13*, 1643180. https://www.frontiersin.org/journals/public-health/articles/10.3389/fpubh.2025.1643180/full

[31] Ferrer, X., van der Vlist, F. N., & Balahur, A. (2024). Fairness and Bias in Artificial Intelligence: A Brief Survey of Sources, Impacts, and Mitigation Strategies. *Information, 6*(1), 3. https://www.mdpi.com/2413-4155/6/1/3

[32] Lo, C. K., & Che, W. S. (2024). Generative AI and Academic Integrity in Higher Education: A Systematic Review and Research Agenda. *Education Sciences, 16*(4), 296. https://www.mdpi.com/2078-2489/16/4/296

[33] Cornell University, Center for Teaching Innovation. (2024). *Ethical AI for Teaching and Learning*. Retrieved September 3, 2025, from https://teaching.cornell.edu/generative-artificial-intelligence/ethical-ai-teaching-and-learning

[34] UNESCO. (2023). *AI competency framework for students*. Retrieved September 3, 2025, from https://www.unesco.org/en/articles/ai-competency-framework-students

[35] AILit. (2025). *Empowering Learners for the Age of AI*. AILit Framework. https://ailiteracyframework.org/wp-content/uploads/2025/05/AILitFramework_ReviewDraft.pdf

[36] Wiley-Blackwell. (n.d.). *British Journal of Educational Technology*. Retrieved September 3, 2025, from https://www.deepdyve.com/browse/journals/0007-1013?rows=10&start=0

[37] Abdmouleh, A., Al-Ahdal, A. A., & Al-Qubati, M. (2024). *The Use of Generative AI Tools in Higher Education: Ethical and Pedagogical Principles*. https://www.researchgate.net/publication/389000405_The_Use_of_Generative_AI_Tools_in_Higher_Education_Ethical_and_Pedagogical_Principles

[38] Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I., & Pechenkina, E. (2024). Beyond efficiency: Empirical insights on generative AI's impact on cognition, metacognition and epistemic agency in learning. *British Journal of Educational Technology*. https://research.monash.edu/en/publications/beyond-efficiency-empirical-insights-on-generative-ais-impact-on

Editor *Changkui LI wtocom@gmail.com*