Peta International Journal of Social Science and Humanity
www.PIJSSh.peta-research.com
Peta Research Publication
Online ISSN 2830-540X

Human-Centered Education in the Age of Artificial Intelligence: A Delegation Framework for Teacher–Student Relationships, Learner Agency, and Ethical Formation

Dai Jinjing[1,*]

[1]Zhejiang Yuexiu University, 312069 Shaoxing, Zhejiang, China
Corresponding Email : 370912256@qq.com

Abstract

As artificial intelligence (AI) systems become increasingly capable of performing instructional, explanatory, assessment, feedback, and tutoring functions, education faces a question that is theoretical as much as technological: which educational functions can appropriately be delegated to AI, and which require continuing human participation because their value depends on relational, agentic, ethical, or accountable forms of educational practice? This article develops the Human-Centered AI Education Framework (HCAI-EF), a theoretical framework for addressing this delegation problem. Drawing on relational pedagogy, theories of subjectification and educational purpose, social-cognitive and self-determination perspectives on agency, and dialogic and care traditions, HCAI-EF conceptualizes human-centered AI-mediated education through three interdependent dimensions: teacher–student relationships, learner agency, and ethical formation. The framework distinguishes three modes of educational delegation: AI-delegable functions, for which automation does not ordinarily alter the central educational purpose of the activity; AI-augmented functions, in which AI may contribute substantially but meaningful human judgment and oversight remain necessary; and human-primary relational functions, for which interpersonal recognition, contextual judgment, moral accountability, or formation constitutes part of the educational value itself. Rather than classifying technologies categorically, the framework evaluates educational functions according to their relational constitutiveness, judgmental complexity, accountability requirements, and consequences for learner agency. It further distinguishes AI replacing educational tasks, augmenting teachers, augmenting learners, and mediating teacher–student interaction, recognizing that AI can expand accessibility, personalization, participation, and autonomy while also creating risks of cognitive substitution, inappropriate reliance, relational displacement, and diminished accountability. The article derives empirically testable propositions and specifies boundary conditions for future research. HCAI-EF therefore shifts the central question from whether AI can perform an educational task to whether, and under what conditions, delegating that task preserves the human purposes through which it acquires educational meaning.

Keywords: artificial intelligence in education; human-centered education; teacher–student relationships; learner agency; ethical formation; AI literacy; relational pedagogy; generative AI

Received on 1 Jun. 2026, Accepted on 15 Jul. 2026, Published on 12 Aug. 2026.

1. Introduction

Artificial intelligence has moved from a specialized educational technology to an increasingly ordinary component of teaching and learning. Adaptive and intelligent tutoring systems can personalize instructional sequences and practice; learning analytics can identify patterns in student activity and performance; and generative AI systems can produce explanations, examples, feedback, assessments, code, and written text on demand (Holmes & Tuomi, 2022; Kasneci et al., 2023; Zawacki-Richter et al., 2019). These developments expand the range of educational functions that can be performed not only *with* computational support but partly *by* computational systems.

The significance of this transformation extends beyond questions of technological effectiveness. As AI becomes capable of participating in activities conventionally associated with teaching explaining, questioning, assessing, advising, providing feedback, and supporting decision-making it changes the division of educational labor among teachers, students, and technological systems. Consequently, the central question is no longer simply whether AI can improve learning outcomes or perform particular instructional tasks effectively. A more fundamental question concerns the criteria by which educational functions should be allocated between humans and AI.

This question matters because education cannot be reduced to the efficient transmission of information or optimization of measurable learning outcomes. Educational traditions concerned with care, dialogue, agency, democratic participation, and subjectification understand education as involving the formation of persons capable of acting, judging, relating to others, and assuming responsibility in a shared world (Biesta, 2010, 2013, 2020; Dewey, 1916; Noddings, 2002). From this perspective, two activities may produce similar immediate outputs while differing substantially in educational meaning. An AI system and a teacher may, for example, provide equally accurate explanations of a concept, yet the pedagogical significance of mentoring a struggling learner, responding to uncertainty, exercising contextual judgment, or engaging a student in ethical deliberation cannot be established by informational accuracy alone.

The emergence of increasingly capable AI therefore requires a distinction among three questions that are often conflated. The first is a capability question: Can an AI system perform a particular educational function? The second is an effectiveness question: Does AI-supported performance of that function produce desirable educational outcomes? The third is a delegation question: Does transferring that function wholly or partly to AI preserve the educational purposes and relationships through which the function acquires its significance? The first two questions are primarily technical and empirical. The third is simultaneously educational, normative, and empirical because it concerns what education is for, which forms of human participation matter, and what consequences follow when those forms of participation are altered.

This article addresses the third question by developing the Human-Centered AI Education Framework (HCAI-EF). The framework does not begin from the assumption that human performance is inherently superior to AI performance. Human teaching can be inconsistent, inaccessible, biased, or constrained by time and institutional conditions, whereas appropriately designed AI can expand access, provide low-stakes practice, support personalization, reduce

routine workload, and enable forms of participation that may otherwise be difficult (Holmes & Tuomi, 2022; Selwyn, 2019). Nor does the framework classify AI technologies as intrinsically beneficial or harmful. Instead, it proposes that human-centeredness should be evaluated according to the educational function being performed, the purposes at stake, the form of human–AI participation, and the consequences of delegation for relationships, learner agency, ethical formation, and accountability.

HCAI-EF makes three related contributions. First, it integrates teacher–student relationships, learner agency, and ethical formation into a common theoretical account of human-centered AI-mediated education. Second, it develops explicit criteria for distinguishing functions that can reasonably be delegated to AI, functions that can be augmented by AI but require meaningful human oversight, and functions for which accountable human participation should remain primary. Third, it translates these theoretical distinctions into empirically examinable propositions and boundary conditions. The resulting framework shifts analysis from the broad question of whether AI belongs in education toward a more discriminating question: under what conditions does AI delegation preserve, strengthen, or undermine the human purposes of educational practice?

2. From AI Capability to Educational Delegation: The Theoretical Problem

AI in education has developed through overlapping technological trajectories. Intelligent tutoring and adaptive systems have a comparatively long history of personalizing sequencing, practice, feedback, and scaffolding (Luckin, 2017; Popenici & Kerr, 2017). Learning analytics and AI-supported assessment have extended computational participation into the interpretation of learner data, prediction of performance and risk, and generation of recommendations (Holmes & Tuomi, 2022). More recently, generative AI has expanded the scope of machine participation from prediction and classification to the production of explanations, dialogue, feedback, written work, and other artifacts central to teaching and assessment (Cotton et al., 2024; Kasneci et al., 2023; Perkins, 2023).

These developments are important not merely because AI systems have become more capable, but because they increasingly participate in activities through which educational relationships and responsibilities are enacted. The relevant transformation is therefore a redistribution of educational functions. An explanation may be produced by a teacher, an AI system, or both. Feedback may be authored by a teacher, generated by AI and reviewed by a teacher, generated directly for a learner, or mediated through an AI-supported communication system. Assessment may similarly involve different configurations of machine inference and human judgment.

The theoretical problem is that **functional equivalence does not necessarily entail educational equivalence**. Education as a practice is partly constituted by the purposes, relationships, judgments, and responsibilities within which particular activities occur (Biesta, 2013; Noddings, 1984; van Manen, 1991). Consequently, determining whether AI can perform an educational function is insufficient for determining whether that function should be delegated to AI.

Consider feedback.

At one level, feedback involves information about the relationship between current and desired performance. At another, feedback may involve interpretation of a learner's history,

confidence, intentions, circumstances, and developmental needs. In still other situations, feedback communicates recognition, expectation, encouragement, disappointment, or moral responsibility. These are not necessarily separate activities; they may be dimensions of the same educational encounter. The degree to which delegation is appropriate therefore depends on what the function requires in a particular context. HCAI-EF conceptualizes this as an educational delegation problem. The central question is not simply:

Can AI perform function F?

but:

Which elements of function F can appropriately be delegated to AI, under what conditions, and what forms of human participation remain necessary to preserve its educational purposes?

This formulation avoids two symmetrical errors. Technological substitutionism assumes that once AI performs a task sufficiently well, human participation becomes unnecessary. Human exceptionalism, by contrast, assumes that human performance is educationally preferable regardless of evidence, context, or function. A human-centered framework requires neither position. It instead asks what form of human participation is educationally consequential and why.

This distinction also means that HCAI-EF classifies functions and function components rather than technologies as such. The same AI system may appropriately perform a largely procedural role in one context, require substantial human oversight in another, and be inappropriate as a substitute for accountable human participation in a third. Automated correction of low-stakes spelling practice, formative feedback on a complex essay, and communication with a learner about academic misconduct may all employ similar underlying technologies while involving very different levels of relationality, judgment, agency, and accountability.

3. Human-Centered Foundations of AI-Mediated Education

The framework draws on four complementary theoretical traditions. Relational pedagogy emphasizes care, pedagogical attentiveness, recognition, and the interpersonal character of educational encounters (Buber, 1970; Noddings, 1984, 2002; van Manen, 1991). Theories of educational purpose and subjectification emphasize that education concerns not only qualification but also how persons come to act, judge, and exist as subjects in relation to others and the world (Biesta, 2010, 2013, 2020). Social-cognitive and self-determination perspectives conceptualize learners as agents whose development depends on intentional action, self-regulation, autonomy, competence, and relatedness (Bandura, 2001; Ryan & Deci, 2000). Dialogic and democratic traditions further locate learning and formation in participation, dialogue, voice, encounter, and shared meaning-making (Dewey, 1916; Freire, 1970; Vygotsky, 1978).

Taken together, these traditions challenge a purely functional conception of education. They suggest that educational quality depends not only on *what output is produced* but also on *how learners participate in producing understanding*, *who exercises judgment and accepts responsibility*, and *what kinds of relationships are enacted through educational activity*. HCAI-EF organizes these concerns around three interconnected dimensions: teacher–student relationships, learner agency, and ethical formation.

3.1 Teacher–Student Relationships and Relationality

Teacher–student relationships are not merely contextual conditions surrounding instruction; they can contribute directly to engagement, belonging, motivation, and achievement. Meta-analytic evidence demonstrates meaningful associations between affective teacher–student relationships and students' engagement and academic outcomes (Roorda et al., 2011), while relational and moral traditions in educational theory emphasize attentiveness, care, expectation, pedagogical tact, and responsibility as constitutive features of teaching (Hansen, 2001; Jackson et al., 1993; Noddings, 1984; van Manen, 1991).

Four constructs are particularly relevant to AI-mediated education. Teacher presence refers here to the learner's perception that the teacher is attentive, available, and responsive. Relational trust concerns confidence in the teacher's benevolence, competence, reliability, and educational commitment. Belonging concerns the learner's experience of recognition and valued membership within an educational community. Pedagogical authority refers to legitimate responsibility for guiding learning and exercising educational judgment rather than authority understood merely as control.

AI can affect these relational conditions in multiple and potentially contradictory ways. Delegating routine activities may enable teachers to devote more time and attention to students, thereby strengthening relational presence. Conversely, efficiency gains may simply be absorbed by additional administrative or instructional demands and therefore produce no relational benefit. AI-mediated communication may enable faster, more accessible, or multilingual interaction, yet it may also obscure whether communication originates from the teacher and reduce cues through which responsiveness and pedagogical attentiveness are perceived.

HCAI-EF therefore treats presence displacement, trust substitution, and mediational thinning as hypothesized mechanisms rather than inevitable consequences of AI use. Presence displacement occurs when AI substitution reduces opportunities for meaningful teacher–student interaction. Trust substitution describes circumstances in which reliance on system output displaces rather than complements confidence in accountable human judgment. Mediational thinning refers to the possibility that technologically mediated interaction reduces relational cues relevant to recognition, responsiveness, or pedagogical tact. Importantly, the opposite effects are also possible: appropriate AI augmentation may expand teachers' capacity to respond and thereby strengthen perceived presence.

The relevant theoretical claim is therefore not that AI cannot generate experiences that learners interpret as supportive, responsive, or trustworthy. Rather, certain educational relationships derive part of their significance from reciprocal human recognition and accountable participation. Whether AI can simulate some phenomenological features of care or presence is consequently different from whether such simulations are educationally equivalent to relationships in which another person can recognize, respond to, and assume responsibility for the learner.

3.2 Learner Agency and Cognitive Delegation

Learner agency concerns the capacity and enacted disposition to influence one's learning through intentional, self-regulated, reflective, and consequential action (Bandura, 2001; Reeve

& Tseng, 2011; Zimmerman, 1990). AI makes agency especially important because it can simultaneously expand what learners are able to do and perform cognitive activities that learners would otherwise undertake themselves.

HCAI-EF distinguishes three analytically related forms of learner agency. Epistemic agency concerns learners' authorship and responsibility in constructing, evaluating, and justifying knowledge. Strategic agency concerns their capacity to plan, monitor, regulate, and adapt learning activity. Decisional agency concerns meaningful influence over the goals, methods, resources, and choices that shape learning. These dimensions make it possible to analyze AI support more precisely than through a general distinction between "empowerment" and "dependence."

AI may strengthen agency by widening access to explanations, enabling learners to seek assistance without social cost, supporting individualized pacing, externalizing ideas for reflection, and providing accessibility functions that increase participation (Holmes & Tuomi, 2022; Selwyn, 2019). Yet the same affordances can weaken agency when AI substitutes for activities through which learners are expected to develop epistemic or regulatory competence.

Cognitive offloading illustrates this ambiguity. Offloading memory, planning, calculation, or search to external tools is a normal feature of human cognition and may improve performance by reducing unnecessary cognitive burden (Risko & Gilbert, 2016). The educational issue is therefore not whether offloading occurs but what is offloaded and what developmental purpose the activity serves. HCAI-EF distinguishes productive offloading, in which AI reduces peripheral cognitive demands while learners retain responsibility for central reasoning, monitoring, and evaluation, from substitutive offloading, in which AI performs the very cognitive activity that learners are expected to develop.

A similar distinction applies to reliance. Frequent AI use should not be equated with diminished agency. Learners may use AI extensively while critically evaluating, revising, and rejecting its outputs. The greater concern is uncritical or substitutive AI reliance: the habitual acceptance or delegation of reasoning without sufficient monitoring, evaluation, or ownership. Research on automation bias and trust in automated systems suggests that inappropriate reliance can reduce vigilance and increase acceptance of erroneous recommendations (Lee & See, 2004; Skitka et al., 1999).

Adaptive systems additionally shape agency through choice architecture. By determining what content, difficulty, recommendation, or activity appears next, such systems may support learners' strategic regulation or quietly replace it (Regan & Jesse, 2019). Human-centered AI design must therefore consider not only whether learners are given choices but whether they retain meaningful authorship over the purposes and processes of their learning.

3.3 Ethical Formation and Critical AI Literacy

AI-mediated education also raises questions that extend beyond learning performance to the development and exercise of ethical judgment. HCAI-EF uses ethical formation to refer specifically to learners' developing capacity to recognize ethically significant situations, deliberate about competing considerations, justify decisions, and act responsibly in AI-mediated educational contexts.

Generative AI makes this particularly visible in questions of authorship and academic integrity. The boundary between legitimate assistance and inappropriate delegation is increasingly difficult to define through simple prohibitions because the educational acceptability of AI use depends on the purpose of the task, the degree of assistance, disclosure, and what competence the assessment is intended to demonstrate (Cotton et al., 2024; Perkins, 2023). Academic integrity under these conditions cannot be treated solely as a detection or compliance problem. It also requires learners to understand why particular forms of delegation preserve or undermine the purposes of learning and assessment.

Ethical formation must nevertheless be distinguished from the broader ethical governance of educational AI. Questions of algorithmic bias, privacy, explainability, institutional accountability, and responsibility for automated decisions concern the conditions under which AI systems are designed and governed (Baker & Hawn, 2022; Gunning et al., 2019; Regan & Jesse, 2019; Santoni de Sio & van den Hoven, 2018). These governance conditions shape learners' experiences, but they are not identical to learners' ethical development. Maintaining this distinction prevents individual students from being made responsible for structural problems that require institutional or regulatory solutions.

Critical AI literacy connects these levels and is treated in HCAI-EF as a cross-cutting enabling competence rather than as a synonym for ethical learning. It includes sufficient understanding of AI capabilities, limitations, uncertainty, data practices, and social consequences to enable informed use and critical evaluation (Long & Magerko, 2020; UNESCO, 2021). Critical AI literacy can therefore influence learner agency, appropriate reliance, integrity judgments, evaluation of AI-generated information, and participation in decisions about educational technologies.

From a human-centered perspective, AI can support ethical learning by presenting cases, alternative perspectives, information, and opportunities for reflection. What should not automatically be delegated is accountability for ethical formation itself. Dialogue about values, responsibility, consequences, and the purposes of education involves forms of judgment and accountability that exceed the provision of ethically relevant information. The educational question is therefore how AI can contribute to ethical deliberation without displacing the human relationships and institutional responsibilities through which learners encounter accountable moral judgment.

4. From Educational Theory to Human–AI Role Configurations

The preceding analysis suggests that AI's educational significance depends not only on technological capability but on the role it occupies within a human–AI educational configuration. Four configurations are particularly useful for analysis: AI replacing educational tasks, AI augmenting teachers, AI augmenting learners, and AI mediating teacher–student interaction.

In task replacement, AI performs an activity that would otherwise be undertaken by a teacher or another human actor. Replacement can be appropriate for bounded and relatively procedural functions, particularly when automation improves availability, consistency, or efficiency without materially changing the educational purpose of the activity. Its risks increase when the

substituted activity requires contextual judgment, relational responsiveness, or accountable decision-making.

In teacher augmentation, AI extends teachers' capacity by supporting planning, information synthesis, formative feedback, analytics, differentiation, or administrative work while teachers retain interpretive and decision-making responsibility. This configuration may be particularly compatible with human-centered education when automation genuinely reallocates human attention toward higher-value pedagogical and relational work. It can nevertheless produce overreliance, professional deskilling, or hidden dependence on biased analytics when oversight becomes nominal rather than meaningful.

In learner augmentation, AI provides explanations, scaffolds, practice, feedback, accessibility support, or cognitive tools directly to students. Such augmentation can increase participation, autonomy, and access, but its educational value depends on whether it expands learners' capacity to act or substitutes for the cognitive and metacognitive activities they are expected to develop.

Finally, in interaction mediation, AI participates in the communicative space between teachers and learners for example, by drafting responses, translating communication, summarizing student questions, personalizing messages, or generating feedback through teacher-facing systems. Mediation is theoretically important because it may either extend teacher responsiveness or obscure authorship and weaken relational transparency.

These configurations should not be interpreted as a hierarchy in which greater human involvement is automatically superior. Human teaching itself can be inaccessible, inconsistent, inequitable, or poorly responsive, whereas AI can create educational opportunities that would otherwise be unavailable. The relevant question is whether a particular configuration supports the educational purposes at stake while preserving appropriate forms of learner agency, human judgment, relational participation, and accountability.

Table 1. Human–AI role configurations in education

Role configuration	Definition	Potential educational opportunities	Principal human-centered risks
AI replacing tasks	AI performs bounded instructional, assessment, feedback, or administrative activities otherwise performed by humans	Availability, consistency, scalability, reduced routine workload	Inappropriate substitution of judgment; reduced human interaction; accountability gaps
AI augmenting teachers	AI supports teacher planning, analysis, feedback, or differentiation while teachers retain responsibility	Greater informational capacity; personalization; potential reallocation of teacher time	Automation bias; deskilling; nominal rather than meaningful oversight; biased analytics
AI augmenting learners	AI provides cognitive, accessibility, practice, or feedback support directly to learners	Accessibility; individualized support; low-stakes practice; expanded learner capacity	Substitutive offloading; uncritical reliance; reduced epistemic or strategic agency

AI mediating interaction	AI generates, transforms, translates, or channels communication between teachers and learners	Timeliness; multilingual access; asynchronous responsiveness; scalable communication	Ambiguous authorship; relational thinning; misplaced trust; misattribution of human attention

5. Research Gap and Framework Requirements

The rapid expansion of human-centered AI scholarship means that the research gap can no longer be defined simply as an absence of attention to ethics, agency, or human control. Rather, three more specific gaps motivate HCAI-EF.

First, there is an educational-purpose gap. Much of the literature on responsible or human-centered AI appropriately emphasizes fairness, transparency, stakeholder participation, trustworthiness, and human control. These principles are necessary, but they do not by themselves determine which forms of human participation are educationally important. A theoretically educational framework must connect decisions about AI to the purposes of education, including learning, relational development, learner agency, subjectification, and ethical formation.

Second, there is a delegation-criteria gap. Discussions of AI adoption often distinguish human and machine roles pragmatically according to technological capability, efficiency, risk, or institutional policy. What remains less explicit is a set of educational criteria for deciding when delegation is appropriate, when augmentation requires meaningful human oversight, and when human participation carries constitutive relational or accountability value.

Third, there is an integrative construct gap. Teacher presence, relational trust, AI reliance, autonomy, critical AI literacy, ethical judgment, and learner agency appear across different research traditions, but their relationships are rarely organized within a common theoretical model capable of generating testable propositions. HCAI-EF addresses this gap by connecting these constructs through an educational account of delegation.

Accordingly, a human-centered delegation framework must satisfy four requirements. It must be purpose-sensitive, evaluating AI use against the educational purposes of the activity rather than technological capability alone; function-specific, because the same technology can occupy different educational roles; context-sensitive, because appropriate delegation changes with developmental level, task stakes, learner vulnerability, and institutional setting; and empirically generative, producing propositions whose observable consequences can be investigated rather than relying exclusively on normative assertion.

6. The Human-Centered AI Education Framework (HCAI-EF)

HCAI-EF integrates three human-centered educational dimensions teacher–student relationships, learner agency, and ethical formation with explicit criteria for determining appropriate modes of AI delegation. Its central premise is that technologies should not be assigned permanently to particular categories. Instead, educational functions and, where necessary, components of functions should be evaluated according to what their successful performance requires.

6.1 Principles and Classification Criteria

Four criteria determine the appropriate degree of AI delegation. Relational constitutiveness concerns the extent to which the educational value of a function depends on interpersonal recognition, reciprocity, presence, or an ongoing pedagogical relationship. A function is highly relationally constitutive when who performs it and the relationship between participants form part of its educational meaning.

Judgmental complexity concerns the extent to which a function requires contextual interpretation rather than bounded procedural inference. Judgmental complexity increases when decisions require integrating ambiguous information about a learner's history, intentions, circumstances, development, or competing educational purposes.

Moral and professional accountability concerns the extent to which an identifiable person must be able to justify, take responsibility for, and respond to the consequences of a decision. Accountability requirements increase with the stakes of the decision and the potential consequences for learners' opportunities, dignity, rights, or development.

Agency consequence concerns the extent to which delegating a function alters learners' opportunities to exercise epistemic, strategic, or decisional agency. Delegation becomes more educationally consequential when AI performs the reasoning, monitoring, choice, or authorship that the learner is expected to develop.

These criteria transform HCAI-EF from a classification of technologies into a contextual decision framework. A single technological capability may therefore occupy different delegation modes in different educational situations. Automated language correction during low-stakes practice, AI-generated formative feedback on an essay, and AI-supported communication about suspected academic misconduct may involve similar generative capabilities but substantially different levels of relationality, judgment, accountability, and agency consequence.

6.2 Three Modes of Educational Delegation

Based on these criteria, HCAI-EF distinguishes three modes rather than three immutable categories of educational activity.

Mode 1: AI-delegable functions

AI-delegable functions are relatively bounded activities for which relational constitutiveness, judgmental complexity, accountability requirements, and agency consequences are ordinarily low. Examples may include routine information retrieval, transcription, formatting, scheduling, administrative logistics, and selected forms of drill and practice. Appropriate delegation can reduce unnecessary workload and potentially release human attention for activities requiring greater pedagogical judgment or relational engagement.

Delegability should nevertheless not be equated with the absence of responsibility. AI outputs may remain inaccurate, biased, or inappropriate, and institutions retain responsibility for selecting and monitoring the systems they deploy. “Minimal oversight” therefore means proportionate oversight, not the absence of human accountability.

Mode 2: AI-augmented functions

AI-augmented functions combine computational capability with consequential pedagogical judgment. Examples include adaptive practice, individualized explanations, formative feedback, writing scaffolds, learning analytics, and assessment support. AI may make substantial contributions to these functions, but teachers or learners must retain meaningful capacity to understand, question, contextualize, reject, or override system outputs when appropriate (Santoni de Sio & van den Hoven, 2018).

The defining principle is meaningful rather than ceremonial human oversight. Human involvement cannot be considered meaningful merely because a teacher technically remains "in the loop." Oversight requires sufficient transparency, competence, time, authority, and responsibility to evaluate AI-supported decisions. The need for such oversight increases with ambiguity, stakes, learner vulnerability, and the consequences of error (Baker & Hawn, 2022; Gunning et al., 2019).

Mode 3: Human-primary relational and accountable functions

The third mode comprises functions for which human participation carries substantial relational, formative, judgmental, or accountability value. Examples include mentoring, sustained expressions of pedagogical care, moral modeling, high-stakes contextual judgments about a learner, dialogue concerning meaning or educational purpose, and sensitive conversations about conduct, identity, difficulty, or development.

HCAI-EF does not claim that AI cannot contribute to these activities. AI may provide information, prompts, translation, accessibility support, or preparatory analysis. The stronger claim is that AI support should not ordinarily displace identifiable and accountable human participation where interpersonal recognition, contextual judgment, or responsibility constitutes part of the educational function itself.

This formulation is deliberately different from declaring such functions technologically "non-delegable." The distinction is educational rather than technical. An AI system may be capable of generating language that resembles encouragement, mentoring, or moral advice. The framework asks instead whether substituting that system for accountable human participation preserves the educational meaning and consequences of the activity.

Some complex educational practices may also span more than one mode. High-stakes assessment, for example, may include AI-delegable administrative processing, AI-augmented analytical interpretation, and human-primary responsibility for consequential judgment and communication. HCAI-EF therefore permits functional decomposition rather than forcing whole educational practices into a single category.

Table 2. HCAI-EF modes of educational delegation

Delegation mode	Typical characteristics	Illustrative functions	Required human involvement	Principal educational purposes at stake

1. AI-delegable	Low relational constitutiveness; bounded judgment; low agency consequence; comparatively low stakes	Routine retrieval; transcription; formatting; scheduling; selected drill and practice	Proportionate monitoring and institutional accountability	Efficiency; accessibility; release of human attention
2. AI-augmented	Moderate-to-high judgmental complexity, agency consequence, or potential impact on learners	Adaptive practice; explanations; formative feedback; writing scaffolds; analytics; scoring support	Meaningful oversight, including transparency, contextualization, challenge, override, and accountable decision-making	Learning quality; fairness; autonomy; epistemic and strategic agency
3. Human-primary relational/accountable	High relational constitutiveness, contextual judgment, moral responsibility, or developmental significance	Mentoring; pedagogical care; moral modeling; sensitive feedback; whole-learner judgment; dialogue about meaning and purpose	Identifiable human participation remains primary; AI may support but should not ordinarily substitute for accountable human involvement	Trust; belonging; subjectification; ethical formation; dignity; responsibility

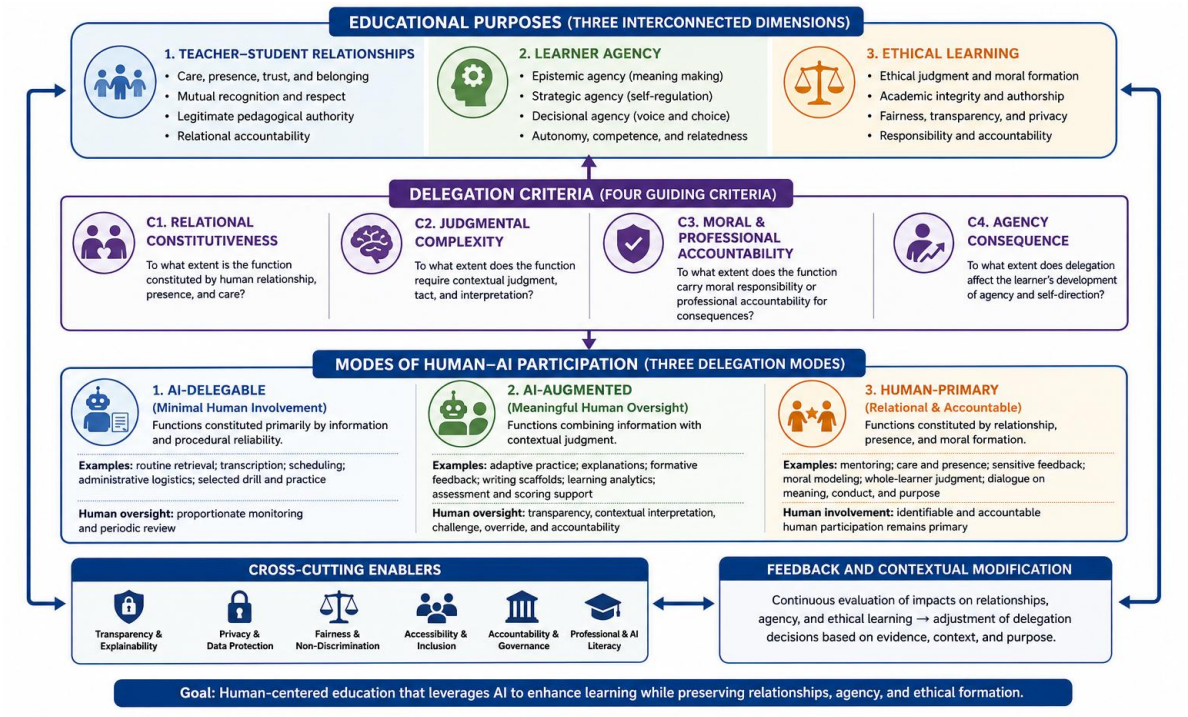

Figure 1. The Human-Centered AI Education Framework (HCAI-EF): educational purposes, delegation criteria, and modes of human–AI participation.

Table 2. HCAI-EF modes of educational delegation, human involvement, and educational purposes

Delegation mode	Illustrative functions	Required human involvement	Principal educational purposes at stake
1. AI-delegable	Routine information retrieval; transcription and formatting; scheduling and administrative logistics; selected forms of drill and practice	Proportionate monitoring and periodic review; institutional accountability remains	Efficiency; accessibility; reduction of routine workload; release of human attention
2. AI-augmented	Adaptive practice; individualized explanations; formative feedback; writing scaffolds; learning analytics; assessment and scoring support	Meaningful human oversight involving transparency, contextual interpretation, challenge, override, and accountability	Learning quality; fairness; learner agency; appropriate personalization
3. Human-primary relational/accountable	Mentoring; pedagogical care and presence; moral modeling; sensitive feedback; whole-learner judgment; dialogue concerning meaning, conduct, and educational purpose	Identifiable and accountable human participation remains primary; AI may support but should not ordinarily substitute for human participation	Trust; belonging; subjectification; ethical formation; dignity; responsibility

7. Theoretical Propositions and Boundary Conditions

HCAI-EF is intended not only as a normative framework but also as an empirically generative one. Its central claims should therefore produce observable implications that can be examined, challenged, and refined. The propositions below concern relationships among modes of AI delegation, human oversight, learner agency, relational experience, critical AI literacy, and educational outcomes. They are deliberately stated as conditional propositions rather than universal predictions because the framework assumes that the consequences of AI use depend on the educational purpose, task characteristics, learner characteristics, institutional arrangements, and form of human participation.

7.1 Relational Reallocation and Teacher Presence

One frequently proposed benefit of educational AI is that automating routine tasks will release teachers' time for higher-value pedagogical activity. HCAI-EF treats this benefit as conditional rather than automatic. Delegating routine functions may reduce workload, but efficiency gains strengthen human-centered education only when the released capacity is actually redirected toward activities involving pedagogical judgment, interaction, mentoring, or responsiveness. Institutions may instead absorb efficiency gains through larger workloads, additional administrative requirements, or increased student-to-teacher ratios.

The relevant mechanism is therefore not automation itself but relational reallocation: the conversion of time or cognitive capacity saved through AI into greater availability for meaningful teacher–student interaction. Because teacher–student relationships are associated

with engagement and achievement (Roorda et al., 2011), this distinction provides a direct empirical link between delegation design and relational outcomes.

P1 Relational reallocation proposition.

Delegation of AI-delegable functions will be associated with stronger perceived teacher presence and relational trust to the extent that the time and attention released by automation are reallocated to meaningful teacher–student interaction. Where such relational reallocation does not occur, delegation alone should not systematically improve relational outcomes.

This proposition improves on a simple "AI saves teachers time" hypothesis because it identifies the organizational mechanism necessary for efficiency to become an educational benefit.

7.2 Human Oversight and Pedagogical Judgment

AI-supported feedback, assessment, and personalization occupy the second delegation mode because they combine computational capabilities with pedagogically consequential interpretation. Meaningful human oversight is therefore expected to matter most when AI outputs are ambiguous, potentially erroneous, consequential for learners, or dependent on contextual knowledge that the system does not adequately represent.

Importantly, HCAI-EF does not assume that human involvement always improves AI-generated outputs. Human reviewers may accept erroneous recommendations, fail to detect bias, or add inconsistent judgments. The theoretical prediction is narrower: human oversight should be beneficial when it is substantive that is, when the human actor possesses sufficient competence, information, time, authority, and willingness to evaluate and override the system rather than merely endorsing its recommendation (Lee & See, 2004; Santoni de Sio & van den Hoven, 2018).

P2 Meaningful-oversight proposition.

For educational functions characterized by substantial judgmental complexity or consequences for learners, AI-supported decisions accompanied by meaningful human oversight will produce more appropriate reliance and fewer consequential uncorrected AI errors than equivalent AI-only arrangements. The magnitude of this advantage will increase with task ambiguity, stakes, and the potential consequences of error.

This proposition deliberately avoids claiming that AI-plus-human feedback must always generate greater learning gains than AI-only feedback. Such an outcome is empirical rather than theoretically guaranteed. Instead, the framework predicts where human judgment should become consequential.

7.3 Cognitive Delegation and Learner Agency

The relationship between AI use and learner agency cannot be captured adequately by frequency of use alone. AI can expand agency by increasing access to information, reducing peripheral cognitive burdens, providing accessibility support, and enabling learners to explore alternatives. It can also diminish agency when it performs cognitive activities that learners themselves need to practice.

The critical distinction is therefore between productive cognitive offloading and substitutive cognitive offloading. Productive offloading transfers peripheral demands to AI while

preserving learner responsibility for central reasoning, monitoring, evaluation, and decision-making. Substitutive offloading transfers the target cognitive process itself. This distinction is especially important in educational settings because short-term task performance and long-term competence development may diverge.

P3 Cognitive delegation proposition.

Substitutive AI offloading will be negatively associated with subsequent independent performance, metacognitive monitoring, and strategic self-regulation when the delegated cognitive process is itself an intended learning objective. Productive offloading of peripheral cognitive demands should show weaker negative associations and may improve performance where it allows learners to devote greater attention to the target learning process.

Critical AI literacy should influence this relationship by improving learners' ability to determine when delegation is appropriate and when independent cognitive engagement is educationally necessary.

P4 Critical-literacy moderation proposition.

The negative relationship between substitutive AI reliance and learner agency will be weaker among learners with higher critical AI literacy because such learners should be better able to evaluate outputs, calibrate reliance, recognize uncertainty, and retain responsibility for consequential reasoning.

Agency also depends on how AI systems structure learner choice. Adaptive recommendations may support autonomy when they provide intelligible options and preserve meaningful learner control, but they may undermine decisional agency when recommendations become opaque prescriptions.

P5 Autonomy-support proposition.

AI-mediated learning environments that provide meaningful choice, transparent recommendations, and opportunities to reject or modify system guidance will be associated with greater perceived autonomy and learner agency than functionally comparable environments in which AI recommendations operate as opaque or effectively mandatory defaults.

This proposition connects HCAI-EF directly to self-determination and agentic perspectives while making the relevant design characteristics observable (Bandura, 2001; Ryan & Deci, 2000).

7.4 AI Literacy, Ethical Formation, and Responsible Use

Critical AI literacy should not be treated simply as technical knowledge about artificial intelligence. Within HCAI-EF, it is a cross-cutting competence that enables learners to recognize system limitations, evaluate outputs, understand the implications of data and automation, and make reasoned decisions about appropriate delegation.

This competence is particularly relevant to academic integrity. Generative AI creates situations in which acceptable and unacceptable assistance cannot always be distinguished through tool use alone. The same technology may support brainstorming, language revision, explanation, or wholesale substitution for assessed work. Responsible use therefore requires learners to

interpret the purpose of the task and determine whether particular forms of AI assistance remain compatible with that purpose (Cotton et al., 2024; Perkins, 2023).

P6 Critical AI literacy and ethical judgment proposition.

Higher critical AI literacy will be associated with better-calibrated AI reliance and more context-sensitive judgments concerning legitimate and illegitimate AI assistance. This relationship should be stronger where expectations regarding authorship, disclosure, and the educational purpose of AI use are explicitly discussed rather than communicated solely through prohibition.

The proposition does not imply that AI literacy alone produces ethical behavior. Institutional norms, incentives, assessment design, peer practices, and perceived fairness may all affect conduct. HCAI-EF therefore treats critical literacy as an enabling condition for ethical judgment rather than a sufficient cause of integrity.

7.5 Equity and Relational Substitution

AI may have particularly important benefits for learners who encounter barriers in conventional educational environments. Translation, alternative modes of representation, individualized pacing, assistive interaction, and low-stakes opportunities for questioning may increase access and participation. At the same time, algorithmic bias, unequal access to high-quality systems, differences in AI literacy, and inappropriate personalization can reproduce or widen existing inequalities (Baker & Hawn, 2022; Regan & Jesse, 2019).

For this reason, HCAI-EF rejects the assumption that AI augmentation is inherently equalizing.

P7 Conditional-equity proposition.

AI augmentation will be more likely to reduce participation or achievement disparities when it removes a clearly identifiable access barrier and is accompanied by equitable access, appropriate human support, and safeguards against biased or inappropriate system outputs. In the absence of these conditions, AI augmentation may leave disparities unchanged or widen them.

The third delegation mode generates a different empirical prediction. When educational activities are strongly relationally constitutive, replacing identifiable human participation with AI may affect perceived care, belonging, trust, or accountability even where informational content is held constant. This is a testable implication of relational pedagogy rather than a claim that students are incapable of experiencing social connection with AI.

P8 Relational-substitution proposition.

For educational functions high in relational constitutiveness such as mentoring, sensitive developmental feedback, or dialogue concerning personal or moral significance AI substitution for accountable human participation will be associated with different levels or forms of perceived care, relational trust, belonging, and accountability than human-primary or human–AI augmented arrangements, even when informational content is comparable.

The direction and magnitude of these differences should be treated as empirical questions and may vary by learner, culture, task, prior relationships, and system design. The proposition therefore permits evidence that some learners experience AI-mediated interactions as

supportive while retaining the theoretical distinction between perceived social responsiveness and reciprocal human accountability.

7.6 Boundary Conditions

The three modes of delegation are not fixed categories that apply identically across learners and settings. Their application depends on several boundary conditions.

Developmental and self-regulatory capacity affects the degree of oversight required. Learners with less developed metacognitive and self-regulatory capacities may require greater scaffolding when using systems capable of performing reasoning, planning, or evaluation on their behalf (Risko & Gilbert, 2016; Zimmerman, 1990). This does not imply that younger learners should necessarily use less AI; rather, it implies that the design of delegation and oversight should reflect developmental capacity.

Task stakes and consequence severity also affect classification. As the consequences of error increase, requirements for accountable human judgment should generally increase. An AI-generated practice recommendation and an AI-supported judgment affecting progression, disciplinary action, or educational opportunity may therefore occupy different delegation modes even when similar computational processes are used.

Learner vulnerability and accessibility needs create a particularly important boundary condition. AI may provide substantial autonomy and access for learners with disabilities or language barriers, meaning that a categorical preference for human delivery could itself reduce agency or inclusion. At the same time, vulnerable learners may face heightened consequences from privacy violations, biased classification, or inappropriate automated decisions (Baker & Hawn, 2022; Regan & Jesse, 2019). Human-centeredness must therefore be evaluated in terms of actual learner capabilities and opportunities rather than the simple amount of human contact.

Educational purpose affects how far delegation can appropriately extend. When an activity primarily serves qualification for example, repetitive practice of a well-defined skill greater automation may be educationally appropriate. Where the activity concerns subjectification, ethical deliberation, identity, belonging, or participation in a shared world, relational and accountability criteria become more salient (Biesta, 2020).

Finally, cultural and institutional context affects interpretations of authority, autonomy, care, teacher responsibility, and acceptable forms of technological mediation. The framework therefore does not assume that a single configuration will be evaluated identically across educational cultures (Noddings, 2002; Schweisfurth, 2013).

These boundary conditions produce a central implication of HCAI-EF: delegation is contextual rather than categorical. The same technological capability may appropriately function as AI-delegable in one situation, AI-augmented in another, and human-primary in a third.

8. Implications for Educational Practice and Governance

The framework has implications at three interconnected levels: teachers' pedagogical practice, learners' use of AI, and institutional governance. These levels should not be treated independently because individual teachers and students cannot sustain human-centered AI practices when organizational incentives, workload arrangements, assessment systems, or technology policies work against them.

8.1 Teachers and Pedagogical Practice

HCAI-EF does not imply that the teacher's role should shift simply from "knowledge provider" to "facilitator," a formulation that understates the complexity of professional teaching. Instead, AI increases the importance of teachers as designers, interpreters, relational professionals, and accountable pedagogical decision-makers.

First, teachers should determine what is being delegated before deciding whether delegation is appropriate. Routine functions can be automated where this improves accessibility or reduces unnecessary workload, but automation should be evaluated partly according to what happens to the capacity it releases. Time saved through AI contributes to human-centered education only when institutional conditions permit that time to improve teaching, interaction, feedback, mentoring, or professional judgment.

Second, teachers require the capacity to exercise meaningful oversight of AI-augmented functions. This involves more than reviewing outputs after they are generated. Teachers need sufficient AI literacy to recognize uncertainty and potential bias, interpret analytics in context, explain the role of AI to learners, challenge recommendations, and assume responsibility for consequential decisions (Gunning et al., 2019; Santoni de Sio & van den Hoven, 2018).

Third, teachers should identify educational encounters in which their own participation carries relational or accountable significance. Sensitive feedback, mentoring, discussions of conduct, responses to learner difficulty, and judgments with substantial consequences should not be transferred to AI merely because systems can generate plausible language. AI may prepare information or support communication, but identifiable human participation should remain visible where responsibility and relationship are part of the educational act.

Finally, teachers themselves should model calibrated AI use. Demonstrating why an AI output is accepted, revised, questioned, or rejected makes critical engagement visible to students and positions AI literacy as part of disciplinary and pedagogical judgment rather than as a separate technical competence.

8.2 Learners and AI-Mediated Agency

For learners, the framework shifts attention from whether AI use is permitted to what forms of AI use preserve and develop agency. Students should learn to distinguish between assistance that expands their capacity and substitution that removes opportunities to develop the competence being assessed or learned.

This distinction can be operationalized through AI-use scaffolds. Before consulting AI, learners may be asked to formulate an initial interpretation, plan, solution, or question. During AI use, they may compare alternatives, request justification, identify uncertainty, or challenge system output. Afterward, they may document what was delegated, what they retained responsibility for, what they changed, and why. Such practices make the boundary between augmentation and substitution pedagogically visible.

Assessment norms should similarly move beyond binary distinctions between "AI allowed" and "AI prohibited." Expectations can specify the kinds of assistance permitted, the educational reason for those boundaries, required disclosure, and which elements of reasoning or

authorship must remain demonstrably the student's own. This approach treats academic integrity as connected to educational purpose rather than merely compliance.

Critical AI literacy should consequently become part of learners' educational entitlement. Students need opportunities to understand AI limitations, uncertainty, bias, data practices, persuasive characteristics, and the ways algorithmic systems structure choice and attention (Long & Magerko, 2020; UNESCO, 2021).

Learner agency also requires procedural voice. Where AI systems materially affect students' learning opportunities or evaluation, students should receive intelligible information about their use and appropriate mechanisms for questioning or contesting consequential outputs.

8.3 Institutional Governance and Policy

Human-centered AI cannot depend exclusively on individual teacher discretion. Institutional arrangements determine which systems are purchased, what data they collect, how workload savings are distributed, what forms of oversight are feasible, and who remains accountable when AI-supported decisions affect students.

Institutions should therefore develop function-specific delegation policies rather than relying solely on lists of approved or prohibited technologies. Such policies should distinguish functions that may ordinarily be automated, functions requiring documented human oversight, and high-stakes or relationally consequential functions for which accountable human participation must remain primary.

Meaningful human oversight also has material requirements. Institutions cannot require teachers to review AI recommendations responsibly while simultaneously increasing workload to a level at which review becomes perfunctory. Governance must therefore address time, professional development, system transparency, authority to override, and clear allocation of responsibility.

Assessment policy requires comparable attention. Institutions should redesign assessment where necessary so that the intended learning remains observable under conditions of widespread generative AI use. Policies should distinguish legitimate augmentation from inappropriate substitution and communicate the educational rationale for these distinctions (Cotton et al., 2024; Perkins, 2023).

Procurement and deployment processes should incorporate privacy, security, accessibility, fairness, and contestability requirements, particularly where systems affect minors, vulnerable learners, or consequential educational decisions (Baker & Hawn, 2022; Regan & Jesse, 2019). Where algorithmic recommendations influence educational opportunity, identifiable human actors and institutions must remain responsible for their interpretation and consequences.

The framework therefore treats accountability as non-transferable: AI may participate in an educational decision, but responsibility for establishing the conditions of use and for consequential institutional decisions remains human and organizational.

9. An Empirical Research Agenda

Testing HCAI-EF requires research designs capable of separating technological capability from delegation structure. Studies that compare “AI use” with “no AI use” are insufficient

because they confound the technology with the way educational responsibility is distributed. Future research should manipulate or measure the mode of delegation, the quality of human oversight, and the educational purpose of the activity.

Experimental studies could compare AI-only, AI-plus-human, and human-primary arrangements while holding informational content as constant as possible. For example, formative feedback experiments could manipulate whether feedback is AI-generated without review, AI-generated and meaningfully reviewed by a teacher, or teacher-generated, and examine learning, error acceptance, trust calibration, perceived presence, and learner agency. Relational-substitution hypotheses could be investigated using mentoring or sensitive-feedback scenarios in which content is standardized while the source and accountability structure vary.

Studies of cognitive delegation should distinguish immediate task performance from subsequent independent performance. Experimental designs could manipulate whether AI supports peripheral processes or performs the target cognitive operation itself and then assess delayed unaided performance, metacognitive calibration, transfer, and self-regulation. Such designs are necessary to test the distinction between productive and substitutive offloading.

Longitudinal studies are particularly important because the educational consequences of repeated AI reliance may not be visible in short-term performance. Tracking learners across semesters could examine whether patterns of AI delegation predict changes in metacognitive monitoring, independent performance, help-seeking, strategic self-regulation, and epistemic agency.

Qualitative and mixed-methods research is equally necessary. Teacher presence, trust, care, accountability, and perceived agency cannot be adequately understood through performance measures alone. Classroom observation, interviews, stimulated recall, and longitudinal case studies could investigate how teachers and learners interpret AI-mediated interactions and when technological mediation is experienced as expanding or diminishing human participation.

Measurement development should accompany this empirical program. Constructs requiring further operationalization include substitutive AI reliance, productive cognitive offloading, teacher presence in AI-mediated environments, relational trust, critical AI literacy, perceived meaningful human oversight, and the three dimensions of epistemic, strategic, and decisional agency. Scale development should follow established procedures for construct definition, item development, validity testing, reliability assessment, and measurement invariance across age and cultural contexts (Boateng et al., 2018; DeVellis, 2017).

Crucially, empirical validation should permit revision of the framework. Evidence that AI-primary relational arrangements produce equivalent or stronger relational outcomes in particular contexts, for example, should not be dismissed because it conflicts with the framework's philosophical starting point. Instead, such evidence should prompt examination of which dimensions of relationship are being measured and whether reciprocal accountability, perceived responsiveness, and educational outcomes behave differently. A theoretically useful framework must expose itself to possible refinement rather than merely organize evidence that confirms its assumptions.

10. Limitations and Scope Conditions

HCAI-EF has several limitations that delimit its claims; first, it is a theoretical framework rather than an empirically validated model. The propositions developed here specify relationships for future testing; they should not be interpreted as established causal effects. Second, the three delegation modes are analytical abstractions. Educational practices frequently contain components with different levels of relationality, judgmental complexity, accountability, and agency consequence. Functional decomposition may therefore be more appropriate than assigning an entire practice to a single mode. Third, the boundaries among modes are necessarily contestable. HCAI-EF provides criteria for reasoned classification rather than an algorithm capable of determining the correct degree of delegation independently of educational judgment. Fourth, the framework is intentionally capability-agnostic. AI systems will continue to change, and future systems may exhibit substantially greater accuracy, personalization, multimodality, and apparent social responsiveness. The framework therefore grounds delegation decisions in educational purposes and consequences rather than in assumptions about current technological limitations. Nevertheless, changes in capability may alter the judgmental or practical conditions under which particular functions are classified. Fifth, the framework draws heavily on relational, agentic, and formative traditions developed largely within Western educational philosophy and psychology. Conceptions of autonomy, authority, care, teacher responsibility, and educational purpose vary across cultures. Cross-cultural examination is therefore necessary rather than optional. Sixth, the framework distinguishes ethical formation from the institutional ethics of AI but does not provide a comprehensive theory of educational AI governance. Issues including political economy, platform power, commercial incentives, environmental costs, labor displacement, intellectual property, and national regulation fall partly outside its immediate scope. These forces may nevertheless determine whether human-centered delegation is institutionally feasible. Finally, some of the framework's foundational claims are normative. Whether reciprocal human recognition or accountable interpersonal participation should be regarded as intrinsically valuable educational goods cannot be settled through empirical evidence alone. Empirical research can, however, test the observable consequences associated with different configurations and clarify where normative assumptions correspond or fail to correspond to learners' experiences and educational outcomes.

11. Conclusion

The rapid development of artificial intelligence makes it increasingly possible to automate activities that have historically been performed by teachers and students. Technological capability, however, does not by itself determine educational desirability. The central question is not simply whether AI can explain, assess, advise, provide feedback, or participate in dialogue, but what happens to the purposes of education when responsibility for these functions is redistributed between humans and machines. The Human-Centered AI Education Framework addresses this problem by linking AI delegation to three interconnected educational dimensions: teacher–student relationships, learner agency, and ethical formation. It proposes four criteria relational constitutiveness, judgmental complexity, moral and professional accountability, and agency consequence for determining whether an educational function is appropriately AI-delegable, AI-augmented, or human-primary. Several principles follow from

this account; purposes should precede technological capability. The appropriate role of AI should be determined by what an educational activity is intended to accomplish, not merely by whether a system can perform it efficiently. Delegation should be function-specific and contextual. Technologies are not inherently delegable or non-delegable. The same AI capability may appropriately occupy different roles depending on the learner, task, stakes, and educational purpose. Human oversight must be meaningful rather than symbolic. Where AI contributes to consequential pedagogical judgment, humans require sufficient understanding, authority, time, and accountability to question and override its outputs. Learner agency should be evaluated developmentally rather than rhetorically. AI supports agency when it expands learners' capacity to understand, decide, regulate, and act; it undermines agency when it substitutes for the very cognitive or metacognitive capacities education is intended to develop. Relational value cannot be inferred from informational performance alone. Where recognition, mentoring, care, trust, or sensitive judgment contributes to the educational meaning of an activity, accountable human participation remains significant even when AI can generate comparable informational content. Ethical AI use requires both learner formation and institutional governance. Students require critical AI literacy and opportunities to develop ethical judgment, while institutions remain responsible for fairness, privacy, transparency, contestability, and the consequences of system deployment. Equity must be evaluated empirically. AI can increase accessibility and participation, but it can also reproduce existing inequalities. Human-centeredness therefore requires attention to actual changes in learners' capabilities and opportunities rather than assumptions that either automation or human delivery is inherently more equitable. Most fundamentally, accountability remains human and institutional. AI systems can contribute information, recommendations, analysis, and interaction, but educational institutions and professionals remain responsible for deciding how those systems are used and for answering to learners when consequential decisions are made. Human-centered education in the age of AI should therefore not be understood as preserving a fixed boundary between humans and machines. Nor should it mean protecting every traditional teacher function from automation. It means designing the distribution of educational activity so that technological augmentation expands human capability without inadvertently displacing the relationships, agency, judgment, and accountability through which education acquires its distinctively human purposes. The question for education is thus no longer simply what AI can do. It is what should be delegated, what should be augmented, what requires accountable human participation, and according to which educational purposes those boundaries should be drawn.

References

Baker, R. S., & Hawn, A. (2022). Algorithmic bias in education. *International Journal of Artificial Intelligence in Education, 32*(4), 1052–1092. https://doi.org/10.1007/s40593-021-00285-9

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology, 52*, 1–26. https://doi.org/10.1146/annurev.psych.52.1.1

Bearman, M., & Ajjawi, R. (2023). Learning to work with the black box: Pedagogy for a world with artificial intelligence. *British Journal of Educational Technology, 54*(5), 1160–1173. https://doi.org/10.1111/bjet.13337

Biesta, G. (2010). *Good education in an age of measurement: Ethics, politics, democracy*. Paradigm Publishers.

Biesta, G. (2013). *The beautiful risk of education*. Paradigm Publishers.

Biesta, G. (2020). Risking ourselves in education: Qualification, socialization, and subjectification revisited. *Educational Theory, 70*(1), 89–104. https://doi.org/10.1111/edth.12411

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health, 6*, Article 149. https://doi.org/10.3389/fpubh.2018.00149

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Buber, M. (1970). *I and thou* (W. Kaufmann, Trans.). Charles Scribner's Sons. (Original work published 1923)

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International, 61*(2), 228–239. https://doi.org/10.1080/14703297.2023.2190148

DeVellis, R. F. (2017). *Scale development: Theory and applications* (4th ed.). SAGE.

Dewey, J. (1916). *Democracy and education: An introduction to the philosophy of education*. Macmillan.

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines, 28*(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5

Freire, P. (1970). *Pedagogy of the oppressed*. Herder and Herder.

Gilligan, C. (1982). *In a different voice: Psychological theory and women's development*. Harvard University Press.

Gunning, D., Stefik, M., Choi, J., Miller, T., Stumpf, S., & Yang, G.-Z. (2019). XAI Explainable artificial intelligence. *Science Robotics, 4*(37), Article eaay7120. https://doi.org/10.1126/scirobotics.aay7120

Hansen, D. T. (2001). *Exploring the moral heart of teaching: Toward a teacher's creed*. Teachers College Press.

Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. *European Journal of Education, 57*(4), 542–570. https://doi.org/10.1111/ejed.12533

Jackson, P. W., Boostrom, R. E., & Hansen, D. T. (1993). *The moral life of schools*. Jossey-Bass.

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence, 1*(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2

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

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors, 46*(1), 50–80. https://doi.org/10.1518/hfes.46.1.50.30392

Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–16). ACM. https://doi.org/10.1145/3313831.3376727

Luckin, R. (2017). Towards artificial intelligence-based assessment systems. *Nature Human Behaviour, 1*, Article 0028. https://doi.org/10.1038/s41562-016-0028

Noddings, N. (1984). *Caring: A feminine approach to ethics and moral education*. University of California Press.

Noddings, N. (2002). *Starting at home: Caring and social policy*. University of California Press.

Perkins, M. (2023). Academic integrity considerations of AI large language models in the post-pandemic era: ChatGPT and beyond. *Journal of University Teaching and Learning Practice, 20*(2). https://doi.org/10.53761/1.20.02.07

Popenici, S. A. D., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research and Practice in Technology Enhanced Learning, 12*, Article 22. https://doi.org/10.1186/s41039-017-0062-8

Reeve, J., & Tseng, C.-M. (2011). Agency as a fourth aspect of students' engagement during learning activities. *Contemporary Educational Psychology, 36*(4), 257–267. https://doi.org/10.1016/j.cedpsych.2011.05.002

Regan, P. M., & Jesse, J. (2019). Ethical challenges of edtech, big data and personalized learning: Twenty-first century student sorting and tracking. *Ethics and Information Technology, 21*(3), 167–179. https://doi.org/10.1007/s10676-018-9492-2

Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences, 20*(9), 676–688. https://doi.org/10.1016/j.tics.2016.07.002

Roorda, D. L., Koomen, H. M. Y., Spilt, J. L., & Oort, F. J. (2011). The influence of affective teacher–student relationships on students' school engagement and achievement: A meta-analytic approach. *Review of Educational Research, 81*(4), 493–529. https://doi.org/10.3102/0034654311421793

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

Santoni de Sio, F., & van den Hoven, J. (2018). Meaningful human control over autonomous systems: A philosophical account. *Frontiers in Robotics and AI, 5*, Article 15. https://doi.org/10.3389/frobt.2018.00015

References (continued)

Schweisfurth, M. (2013). *Learner-centred education in international perspective: Whose pedagogy for whose development?* Routledge.

Selwyn, N. (2019). *Should robots replace teachers? AI and the future of education*. Polity Press.

Selwyn, N. (2022). The future of AI and education: Some cautionary notes. *European Journal of Education, 57*(4), 620–631. https://doi.org/10.1111/ejed.12532

Skitka, L. J., Mosier, K. L., & Burdick, M. (1999). Does automation bias decision-making? *International Journal of Human-Computer Studies, 51*(5), 991–1006. https://doi.org/10.1006/ijhc.1999.0252

UNESCO. (2021). *AI and education: Guidance for policy-makers*. UNESCO.

van Manen, M. (1991). *The tact of teaching: The meaning of pedagogical thoughtfulness*. State University of New York Press.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Williamson, B., & Eynon, R. (2020). Historical threads, missing links, and future directions in AI in education. *Learning, Media and Technology, 45*(3), 223–235. https://doi.org/10.1080/17439884.2020.1798995

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education Where are the educators? *International Journal of Educational Technology in Higher Education, 16*, Article 39. https://doi.org/10.1186/s41239-019-0171-0

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist, 25*(1), 3–17. https://doi.org/10.1207/s15326985ep2501_2