Social Science Chronicle

OPEN ACCESS

Peer-Reviewed, International, Academic Research Journal

ISSN : 3048-6297

Citation
Guberina, T., & Procházka, F. (2026). Human-AI Collaborative Teaching: Generative Artificial Intelligence (Gen-AI) as Co-Teacher. *Social Science Chronicle*, Vol. 6, Issue - 1, pp. 1-19.

Digital Object Identifier (DOI)
https://doi.org/10.56106/ssc.2026.003

Received - August 18, 2025
Accepted - January 25, 2026
Published - Online First

Web-Link
All contents of this peer-reviewed article, along with other relevant details are available at:
http://socialsciencechronicle.com/article-ssc-2026-003

RESEARCH ARTICLE

Human-AI Collaborative Teaching: Generative Artificial Intelligence (Gen-AI) as Co-Teacher

Tajana Guberina[1]* and Filip Procházka[2]

[1] FH Joanneum University of Applied Sciences, Graz, Austria.
[2] University of South Bohemia, České Budějovice, Czech Republic.

* *Corresponding Author*
E-mail : guberinatajana@yahoo.com

Abstract

This article theorizes Human-AI Collaborative Teaching as a co-teacher paradigm grounded in joint cognitive systems and reliability-first sociotechnical design, where instructional quality emerges from coupling, constraint and accountable orchestration rather than model fluency. The synthesis reframes teaching as real-time regulation under bounded rationality, linking distributed cognition and situated cognition to role-bearing AI participation across planning, enactment, assessment and reflection. It specifies a governance-ready architecture in which teacher authority is preserved through decision-rights partitioning, mixed-initiative interaction protocols, calibrated uncertainty signalling, and an abstain-escalate safety regime. Epistemic robustness is operationalized through provenance discipline, evidence-anchored feedback and contestability pathways that protect epistemic dignity, participation equity and multilingual-accessibility rights under high-stakes accountability. The article integrates constructs from learning sciences, human-computer interaction, resilience engineering, implementation science and public governance to produce five compact design instruments, a theoretical lens map, role-ecology contracts, interaction protocol patterns, a governance risk register, and an institutional maturity model for scalable adoption. The resulting framework offers concrete, globally portable implementation logic for policy makers, workforce development leaders, and educational technologists seeking audit-ready co-teaching infrastructures that enhance teacher noticing, strengthen formative inference and sustain assessment validity without privacy erosion, surveillance creep or de-skilling.

Keywords
Human-AI Collaboration, AI in Education, Educational Technology, Co-Teaching, Hybrid Intelligence, Human-AI Interaction, Distributed Cognition, Joint Cognitive Systems, Learning Analytics, Instructional Orchestration, Trust Calibration, Educational Data Governance, Academic Integrity.

1. Introduction

Teaching is a high-velocity, high-stakes control activity in which a professional must continuously infer learner states, regulate task difficulty, and sustain epistemic momentum while managing attention, affect, and social order. The co-teacher paradigm positions artificial intelligence not as a peripheral productivity aid but as a role-bearing partner that participates in instructional sensemaking, decision-shaping, and adaptive orchestration across the lesson cycle (Hutson, 2024; Groothuijsen et al., 2024; Sajja et al., 2024). This shift is enabled by large-scale language and multimodal systems, retrieval-grounded generation, and agentic tool-use that can draft explanations, propose probes, simulate misconceptions, and coordinate workflow sequences. Yet classroom viability is not a question of fluency, it is a question of reliability under constraint, including latency, privacy, integrity, equity, and institutional accountability (Gupta, 2025; Wu et al., 2024; Lata, 2024). A co-teacher must therefore be theorized as an instructional subsystem with explicit decision rights, verification routines, and recovery paths, aligned to global governance expectations and workforce credentialing pressures for safe scaling across multilingual, low-resource, and high-stakes learning ecologies.

Definitions, Scope Boundaries, and Construct Differentiation

An AI co-teacher is defined here as a sociotechnical system that occupies an explicit instructional role, influences pedagogical choices, and operates under teacher-governed authority allocation across planning, enactment, assessment, and reflection. The construct requires role explicitness, bounded permissions, and traceable handoffs, otherwise the system is an assistant, not a co-teacher (Zeb et al., 2025; Jacques et al., 2024; Nurhasanah & Nugraha, 2024). It differs from a student-facing tutor that optimizes individualized practice without whole-class orchestration, and it differs from analytics dashboards that summarize signals without participating in instructional dialogue. It also differs from automated grading engines because co-teaching demands criteria-based interpretation, contestability, and pedagogical follow-through, not only scoring (Al-Kfairy et al., 2024; Melweth et al., 2024; Qian & Wexler, 2024). Scope boundaries are therefore set around formal education settings where professional accountability is non-delegable, and around workflows in which decision rights, verification discipline, provenance control, and abstain-escalate protocols can be operationalized to protect instructional integrity. The review treats K-12 and higher education as distinct governance regimes, distinguishes teacher-facing backchannel support from student-facing interaction, and assumes multilingual, accessibility-sensitive classrooms as a baseline rather than an edge case.

Stance and Integrative Synthesis Logic

This narrative review is conceptual-theoretical rather than empirical, its goal is to build an explanatory architecture that makes co-teaching designable, governable, and evaluable as a coherent instructional system. The synthesis proceeds by mechanism-based reasoning that integrates constructs from bounded rationality, distributed cognition, situated cognition, joint cognitive systems, activity theory, sociomateriality, cognitive load theory, ICAP, dialogic pedagogy, epistemic cognition, trust calibration, mixed-initiative interaction, and resilience engineering. Across these lenses, the paper treats instruction as continuous regulation, collaboration as coupling quality, and reliability as the joint product of interaction protocols and governance constraints. The analytic method is to specify necessary features and failure modes, then derive design primitives, including control points, uncertainty signalling, provenance discipline, contestability pathways, and recovery-from-error routines. This approach prioritizes construct clarity and institutional feasibility over technological spectacle, so that the co-teacher paradigm can be implemented without undermining professional judgment or educational legitimacy. It also integrates normative constraints from value-sensitive design and rights-based policy logics.

Intended Contributions for Research, Policy, and Practice

The article contributes a multi-layer conceptual model of human-AI collaborative teaching that links cognition, interaction, role ecology, and governance into a single analytic frame. At the cognition layer, it specifies how a co-teacher can externalize teacher noticing, compress planning overhead, and increase the precision of formative inference without displacing professional judgment, through constrained generation anchored to curriculum criteria and evidence of learning. At the role-ecology layer, it formalizes division-of-labour patterns, separating suggestion authority, decision authority, and execution authority, thereby preventing hidden delegation and deskilling. At the interaction layer, it defines orchestration regimes that regulate initiative, timing, repair, and participation equity, making classroom flow a first-class design requirement. At the governance layer, it operationalizes audit-ready accountability, data minimization, integrity protection, and student contestability as instructional infrastructure rather than compliance overhead. The five tables in Sections 2 through 6 function as compact design instruments that translate theory into implementable specifications.

Roadmap of the Seven-Section Argument

Section 2 develops the theoretical foundations that make the co-teacher paradigm intelligible as a joint cognitive and sociotechnical arrangement, and Table 1 consolidates the major theoretical lenses with their design implications. Section 3 constructs a role ecology and division-of-instructional-labour model, and Table 2 specifies authority, responsibility, and escalation conditions across co-teacher roles. Section 4 formalizes interaction architectures and orchestration logics, and Table 3 enumerates protocol patterns that preserve classroom flow while controlling epistemic risk. Section 5 treats epistemic governance, ethics, and accountability as enabling infrastructure, and Table 4 maps risk domains to concrete control primitives and ownership boundaries. Section 6 translates the model into design principles and institutional implementation logic, and Table 5 provides an adoption maturity model that aligns technical configuration with professional learning, assessment validity, and auditability.

Section 7 synthesizes the argument into actionable conceptual commitments for global educational systems. Each table is positioned early within its section to support cross-sectional coherence.

2. Theoretical Foundations for AI as a Co-Teacher

In formal learning environments, teaching is best modeled as constrained cognition under bounded rationality, where a professional continuously satisficies among competing pedagogical goods, including conceptual rigor, time-on-task, affect regulation, and participation equity (Wang et al., 2025; Kaspersen et al., 2024). The co-teacher paradigm becomes intelligible when instruction is reframed as a dynamic regulation problem, not a content transmission problem, because teachers operate with partial observability, noisy signals, and severe opportunity costs. A viable AI co-teacher must therefore function as a cognitive prosthesis that compresses search, stabilizes pacing decisions, and augments diagnostic resolution without annexing professional discretion (Bozkurt et al., 2024; Crawford et al., 2024; Park & Ahn, 2024). This requires mechanistic commitments to ecological rationality, where the quality of a decision is evaluated by its context-fit rather than its abstract optimality. The theory-to-design translations that follow, including decision-rights partitioning, uncertainty articulation, and recovery loops, are consolidated in Table 1 for audit-ready implementation.

Distributed and Situated Cognition in Sociomaterial Learning Ecologies

The co-teacher construct is further grounded in distributed cognition, where instructional intelligence is not located in any single agent but is distributed across teachers, learners, artifacts, and infrastructures such as rubrics, exemplars, curricular maps, and platform logs. When co-teaching is treated as distributed work, AI outputs cease to be authoritative answers and become boundary objects that must be integrated, contested, and contextualized within classroom activity (Yeter et al., 2024; Chuang, 2024; Asad & Ajaz, 2024). Situated cognition adds a critical constraint that meanings, misconceptions, and motivational states are emergent properties of participation structures, linguistic repertoires, and local norms, rather than static attributes of individuals. A co-teacher must therefore support context-sensitive sensemaking, for instance by mapping prompts to task demands, discourse norms, and accessibility constraints, while avoiding decontextualized fluency that can intensify misunderstanding (Fissore et al., 2024; Pratschke, 2024; Kolbjørnsrud, 2024). This lens also implies representational discipline, where the co-teacher must preserve curricular intent and epistemic standards in every transformation of instructional material.

Joint Cognitive Systems and Reliability-First Team Cognition

The most stringent conceptual requirement for AI co-teaching is the joint cognitive systems framing, where reliability emerges from coupling quality, not from component capability. In this view, the teacher and AI form a coordinated cognitive assemblage whose performance depends on shared situation awareness, common ground, and low-friction coordination under classroom time pressure (Li, H., et al., 2024; Przegalinska & Triantoro, 2024; Jung & Suh, 2024). Team cognition constructs, including shared mental models and transactive memory, imply that the co-teacher must be predictable in what it can do, explicit in what it cannot do, and disciplined in when it should abstain. Resilience engineering extends the team view by elevating recovery-from-error to a first-class design criterion, because classrooms are adversarial in the benign sense, they contain ambiguity, novelty, and social complexity that produce inevitable breakdowns. Accordingly, Table 1 specifies coupling primitives such as escalation thresholds, verification checkpoints, and correction workflows that prevent silent error propagation.

Instructional Constraint Theories and Designable Co-Teacher Primitives

Instructional legitimacy requires that the co-teacher be constrained by learning-science constructs that govern cognitive efficiency and epistemic quality, including cognitive load theory, cognitive apprenticeship, ICAP, and criteria-referenced formative assessment. These constructs impose non-negotiable requirements, such as minimizing split-attention burdens, sustaining generative activity rather than answer-first dependency, and linking feedback to action and task criteria rather than stylistic persuasion (Joseph & Uzondu, 2024; Fu & Weng, 2024; Ruiz-Rojas et al., 2024). In addition, sociotechnical and normative lenses, including activity theory, sociomateriality, critical pedagogy, and value-sensitive design, require explicit governance over division of labour, authority allocation, and contestability, because co-teaching reorganizes power and voice in the classroom. Table 1 synthesizes these theoretical constraints into implementable primitives that can be embedded in interaction protocols and institutional policies, ensuring that "co-teaching" denotes a governed instructional system rather than an unbounded conversational convenience.

Table 1. Theoretical Lenses and Design Implications for AI Co-Teaching

Lens Anchor and Epistemic Frame	Instructional Problem Reframing	Core Constructs and Mechanistic Commitments	Designable Co-Teacher Primitives	Predictable Failure Mode Without the Lens
Bounded and Ecological Rationality	*Teaching as satisficing under scarcity, attention economy, and uncertainty*	*Bounded rationality, ecological rationality, decision heuristics, constraint satisfaction*	*Decision-rights partitioning, time-bounded suggestioning, uncertainty marking, teacher override by default*	*Overcomplex advice streams, brittle optimization, classroom flow disruption*

Distributed and Situated Cognition	*Instructional intelligence distributed across people, artifacts, and context*	*Distributed cognition, situated cognition, boundary objects, context dependence*	*Artifact-centered outputs, curriculum-grounded transformations, context prompts, locale-aware language registers*	*Decontextualized fluency, misaligned examples, culturally discordant explanations*
Joint Cognitive Systems and Team Cognition	*Reliability as coupling quality rather than component accuracy*	*Joint cognitive systems, shared situation awareness, shared mental models, coordination costs*	*Escalation thresholds, verification checkpoints, predictable capability disclosure, repair workflows*	*Silent error propagation, authority conflict, unstable coordination under time pressure*
Instructional Design and Cognitive Efficiency	*Learning as regulated engagement, not information exposure*	*Cognitive load theory, ICAP, cognitive apprenticeship, generative processing*	*Cognitive load budgeting, stepwise scaffolding and fading, activity prompts that demand justification, worked-example governance*	*Illusion of comprehension, answer-first dependency, superficial engagement masking fragility*
Formative Inference and Evidence-Centered Assessment	*Feedback as inference tied to criteria, evidence, and action*	*Criteria-referenced judgment, evidence models, actionable feedback loops*	*Rubric grounding, evidence citation from student work, feedback-to-next-step templates, abstain on missing evidence*	*Verbose feedback without actionability, hallucinated claims about learner performance*
Human-AI Interaction and Governance-First Sociotechnics	*Co-teaching as power-aware infrastructure with accountability*	*Trust calibration, automation bias, mixed-initiative interaction, value-sensitive design, contestability*	*Calibrated confidence cues, abstain-escalate protocol, provenance discipline, audit logging, appeal pathways*	*Overreliance, surveillance creep, deskilling, inequitable error distribution, legitimacy collapse*

Table 1 is intentionally design-forward, because co-teaching is a high-accountability practice where conceptual elegance must resolve into operational controls. Notably, the table's primitives converge on a small set of reliability mechanisms, including calibrated uncertainty, bounded initiative, explicit handoffs, evidence discipline, and recovery routines, which collectively transform generic model fluency into instructionally trustworthy participation. This consolidation also clarifies why "AI accuracy" is an insufficient metric, since pedagogical harm can arise from timing, framing, power effects, and misalignment even when outputs are superficially plausible. The next subsections extend this foundation by specifying how sociocultural legitimacy and human-AI interaction dynamics translate these primitives into classroom-viable collaboration regimes, while preserving teacher agency and student epistemic dignity.

Socio-Ccultural Legitimacy, Epistemic Justice, and Authority Allocation

Co-teaching is also a legitimacy problem because instruction is embedded in social norms that govern who can speak, what counts as knowledge, and whose reasoning is recognized as credible. Dialogic pedagogy and epistemic discourse norms imply that a co-teacher must support accountable talk, justification, counterexample testing, and principled disagreement, rather than collapsing inquiry into premature closure (Noroozi et al., 2024; Milana et al., 2024; Kim et al., 2025). Equity constraints, informed by constructs of epistemic justice and culturally sustaining practice, require that co-teacher mediation not privilege dominant language varieties, background assumptions, or normative exemplars that marginalize learners' lived knowledge. Activity theory further emphasizes that contradictions emerge when tools reconfigure division of labour, for instance when AI feedback implicitly becomes evaluation, thereby shifting authority away from transparent criteria and teacher judgment. For this reason, Table 1 foregrounds contestability, provenance discipline, and decision-rights partitioning as legitimacy-preserving controls that make co-teaching accountable to classroom norms and institutional obligations.

Human-AI Interaction Dynamics for Trustworthy Pedagogical Coupling

Finally, co-teaching depends on human-AI interaction dynamics that calibrate reliance in proportion to risk, observability, and reversibility. Trust calibration requires the co-teacher to signal uncertainty in an interpretable way, enabling appropriate reliance while reducing automation bias and compliance drift (Laak & Aru, 2025; Storozhyk, 2024; Tariq, 2025). Mixed-initiative interaction provides a principled language for allocating initiative across teacher-led, AI-suggested, and bounded-autonomous micro-actions, ensuring that timing and turn-taking align with the classroom's interaction order. The co-teacher must also support conversational grounding and repair, including clarification prompts, assumption checks, and fast correction pathways, because pedagogical harm is often produced by uncorrected misinterpretation rather than explicit misinformation. These interaction commitments complete the theoretical architecture by converting abstract constraints into usable collaboration regimes, and they connect directly to the design primitives consolidated in Table 1, which will be operationalized as role ecologies in Section 3 and orchestration protocols in Section 4.

3. Role Ecologies and Division of Instructional Labour

Human-AI co-teaching becomes operationally meaningful only when it is treated as a role ecology, not a feature bundle, because classrooms are interaction-dense institutions where authority, legitimacy, and accountability are continuously negotiated under time pressure (Mao et al., 2024; Cai et al., 2025; Cingillioglu et al., 2024). A role ecology specifies how instructional agency is distributed across teacher, learners, AI, curricular artifacts, and institutional rules, thereby preventing role drift and covert delegation that can erode professional discretion. In this ecology, the AI co-teacher is not an autonomous actor, it is a governed subsystem whose participation must be bounded by decision rights, verification discipline, and recovery competence, consistent with the coupling primitives articulated in Table 1. The practical implication is that every co-teacher capability must be paired with an explicit locus of control and a known escalation route, because otherwise "help" becomes indistinguishable from unaccountable influence. Table 2 will formalize this role ecology into compact authority and responsibility regimes that can be enacted across instructional contexts.

Division of Labour Across the Instructional Cycle Under Institutional Constraints

Division of instructional labour is best modeled as a tri-phasic cycle, pre-active planning, interactive enactment, and post-active assessment-reflection, each with distinct cognitive demands, latency constraints, and error surfaces. In the pre-active phase, co-teaching value emerges from compressing search and synthesis while preserving curricular alignment, epistemic rigor, and accessibility constraints, because planning is a constraint-satisfaction problem spanning time, resources, learner variability, and assessment coherence (Yadav, 2025; Goldman et al., 2024). In the interactive phase, the dominant problem is orchestration under partial observability, where the teacher must regulate pacing, discourse, and participation equity while diagnosing misconceptions in real time, a setting where unverified AI fluency can destabilize classroom order. In the post-active phase, the critical work is evidence-centered interpretation, feedback actionability, and re-teaching decisions, where the AI must be constrained to rubric-grounded inference rather than narrative persuasion. These phase-specific demands motivate differentiated authority regimes and output envelopes, which are specified as implementable role contracts in Table 2.

Pedagogically Legitimate Co-Teacher Roles and Instructional Function Differentiation

A coherent co-teacher taxonomy must be functionally differentiated by instructional intent, not by interface aesthetics, because pedagogical legitimacy depends on whether the role strengthens learning mechanisms rather than simply accelerating production. The instructional architect role formalizes objective-activity-evidence coherence, translating curriculum into task sequences that preserve epistemic demand while enabling differentiation (Peters & Green, 2024; Li, Z., et al., 2024; Leong & Zhang, 2025). The misconception diagnostician role is legitimate only when it operates as hypothesis generation with explicit uncertainty, because diagnosis without evidentiary discipline can harden deficit framings. The dialogic facilitator role supports dialogic pedagogy by generating talk-moves that elicit justification, counterexample testing, and conceptual repair, without scripting inauthentic discourse (Kim, T., et al., 2024; Han et al., 2024; Bozkurt, 2024). The formative assessment partner role must be governed by evidence-centered logic, where claims about learning are anchored to observable artifacts and criteria. The inclusion and accessibility partner role operationalizes universal design for learning through representational variation and barrier removal while preserving rigor. These roles require distinct decision rights, verification gates, and escalation triggers, consolidated into a compact governance-ready schema in Table 2.

Authority Allocation, Professional Agency, and Audit-Ready Responsibility Regimes

Professional agency is the normative anchor of co-teaching because teaching is a high-accountability practice in which discretion is inseparable from ethical obligation, institutional mandate, and learner safeguarding (Yim & Su, 2025; Guggemos, 2024; Anane-Simon & Atiku, 2024). Authority allocation must therefore separate suggestion authority from decision authority and execution authority, with the teacher retaining final dispositional control as the default configuration, and any bounded delegation treated as an explicit policy choice rather than a convenience setting. This is not merely philosophical, it is an operational requirement for preventing automation bias, deskilling trajectories, and silent responsibility transfer (Tarisayi, 2024; Sowa & Przegalinska, 2025; Jiang et al., 2024). Audit-ready responsibility regimes also require that co-teacher actions be legible, reversible, and attributable, including records of overrides, abstentions, and escalation events, consistent with reliability-first joint cognition. Table 2 specifies role archetypes, instructional loci, authority and verification regimes, input-output envelopes, and risk-mitigation primitives, enabling institutions to configure co-teaching in a way that remains pedagogically coherent and governance-compliant.

Table 2. Role Ecology Contracts for Accountable Human-AI Co-Teaching

Role Archetype Identifier	Instructional Locus in the Lesson Cycle	Authority and Verification Regime	Input-Output Envelope Specification	Risk Profile and Mitigation Primitive
Instructional Architect	*Pre-active planning, sequencing, alignment*	*Teacher decides, AI proposes, verification required at objective-activity-evidence junctions*	*Inputs curriculum map, constraints, learner variability signals, outputs lesson flow, task variants, alignment checks*	*Alignment drift mitigated by provenance discipline, constraint visibility, override logging*
Misconception Diagnostician	*Interactive enactment, targeted probing, conceptual repair*	*Teacher approves probes, AI marks uncertainty, escalation on low-evidence states*	*Inputs learner responses and task criteria, outputs hypothesis set, probe prompts, error signatures*	*Deficit labeling mitigated by uncertainty articulation, evidence thresholds, contestability cues*
Dialogic Discourse Facilitator	*Whole-class discussion, small-group talk, participation equity*	*Teacher moderates, AI suggests talk-moves, verification via discourse intent alignment*	*Inputs discussion goal, discourse norms, outputs revoicing prompts, justification cues, counterexample prompts*	*Scripted pseudo-dialogue mitigated by authenticity constraints, teacher personalization, abstain triggers*
Formative Assessment Partner	*Post-active interpretation, feedback, re-teaching decisions*	*Teacher validates judgments, AI drafts criteria-referenced feedback, mandatory rubric grounding*	*Inputs rubric and student work artifacts, outputs feedback drafts, next-step actions, misconception links*	*Hallucinated claims mitigated by evidence anchoring, abstain-escalate protocol, audit trail*
Inclusion and Accessibility Adapter	*Pre-active material design, interactive supports, barrier removal*	*Teacher authorizes accommodations, AI transforms formats under rigor-preservation checks*	*Inputs accessibility needs, language levels, outputs alternative representations, scaffolded supports*	*Oversimplification mitigated by epistemic demand checks, accessibility compliance routines, review gates*
Orchestration and Pacing Aide	*Interactive classroom flow, grouping, transitions, timeboxing*	*Teacher retains execution control, AI flags signals, escalation on ambiguity or privacy risk*	*Inputs time constraints and activity state, outputs pacing cues, grouping suggestions, transition prompts*	*Stigmatizing grouping mitigated by fairness constraints, explainable rationale, teacher override defaults*

Table 2 functions as a configurational blueprint that translates role ecology into enforceable instructional contracts, thereby preventing the common failure in which AI participation expands opportunistically without corresponding accountability scaffolds. The table also clarifies that co-teaching is not a monolithic capability, it is a portfolio of bounded roles, each with its own verification discipline and failure modes, which is essential for reliability under classroom constraints. By specifying input-output envelopes, the table implicitly enforces data minimization and purpose limitation, since roles that do not require sensitive learner data should be configured without it. The next subsection extends this role ecology by situating student agency within a triadic classroom contract, ensuring that co-teaching does not collapse authorship, self-regulation, or epistemic dignity.

Student Agency, Triadic Classroom Contract, and Epistemic Dignity Preservation

A co-teaching ecology is incomplete without an explicit account of student agency, because the teacher-student-AI triad reshapes authorship norms, help-seeking behavior, and epistemic responsibility. The triadic contract requires that students understand what kinds of assistance are legitimate, how AI-mediated feedback should be interpreted, and how to contest outputs that misrepresent their intent, competence, or cultural knowledge (Rane & Choudhary, 2024; Yadav & Shrawankar, 2025; Łodzikowski et al., 2024). From a self-regulated learning lens, co-teaching should scaffold planning, monitoring, and reflection while preserving productive struggle, otherwise the system becomes a cognitive outsourcing channel that degrades metacognitive calibration. Integrity is therefore a designable norm, not a policing afterthought, and Table 2 supports this by constraining assessment-adjacent roles to evidence-anchored drafts subject to teacher validation. Epistemic dignity also demands that co-teaching preserve multilingual repertoires and accessibility needs without stigmatization, which requires contestability pathways and teacher-governed transparency in how instructional decisions are shaped by AI suggestions.

4. Interaction Architectures and Orchestration Logics

Human-AI co-teaching becomes classroom-viable only when interaction is engineered as infrastructure, because instructional time is a scarce resource and the classroom is a latency-intolerant, attention-fragile, socially complex setting. An interaction architecture specifies when the co-teacher can speak, what it can see, what it can store, what it can propose, and how its proposals can be verified, rejected, or repaired without destabilizing flow. This requires treating classroom interaction as an economy of interruptions, where each AI intervention imposes cognitive switching costs, discourse reorientation costs, and credibility risks (Liang & Bai, 2025; Chee et al., 2025; Adel & HS Alani, 2024). Accordingly, the co-teacher must operate through friction-minimizing micro-interventions, bounded suggestions, and teacher-governed control points, consistent with the coupling primitives summarized earlier in Table 1 and operational role contracts in Table 2. A robust architecture also internalizes resilience engineering by assuming inevitable breakdowns, thereby prioritizing repair, rollback, and traceable correction over performative fluency. Table 3 will codify these interaction patterns into protocol specifications that are auditable, teachable, and portable across global classrooms.

Temporal Orchestration, Multi-Scale Control Loops, and Stability Conditions

Instructional orchestration unfolds across nested temporal scales, micro loops that operate within seconds and minutes, meso loops that shape whole-lesson phase transitions, and macro loops that regulate unit pacing, cumulative evidence, and curricular progression. A co-teacher that ignores temporal structure risks overcorrecting, creating oscillations in pacing and cognitive load, or undercorrecting, allowing misconceptions and disengagement to accumulate (Mustafa et al., 2024; Fan et al., 2025; Peláez-Sánchez et al., 2024). A control-theoretic framing is therefore conceptually useful, where classroom signals are partial and noisy, interventions have delayed effects, and stability depends on calibrated gain and carefully set thresholds. Micro-scale interventions include question prompts, revoicing candidates, and immediate checks for understanding, while meso-scale interventions include timeboxing cues, group reconfiguration suggestions, and transition scripts that preserve epistemic continuity (Yu et al., 2025; Yan, L., Martinez-Maldonado, & Gasevic, 2024). Macro-scale orchestration includes spiral review planning, interleaving schedules, and assessment readiness sequencing, all governed by constraint satisfaction under institutional calendars. These temporal considerations anticipate the need for explicit trigger conditions, initiative allocation, and abstain-escalate regimes, which Table 3 will articulate as portable protocol primitives rather than ad hoc conversational habits.

Mixed-Initiative Regimes, Turn-Taking Protocols, and Interactional Governance

A co-teacher must be designed for mixed-initiative interaction because neither pure teacher-led querying nor continuous AI proactivity is stable in real classrooms. Teacher-led advisory regimes preserve authority but can underutilize AI sensing and synthesis, whereas AI-led regimes amplify automation bias, create voice displacement, and increase the risk of illegible influence (Karakose & Tülübas, 2024; Suryanarayana et al., 2024; Imran et al., 2024). Mixed-initiative designs therefore require explicit initiative contracts that specify what cues can trigger AI suggestions, what suggestion types are permissible, and what verification steps must precede any instructional enactment. Turn-taking protocols must be compatible with classroom discourse norms, including the teacher's right to withhold, defer, or redirect AI contributions to preserve coherence and relational trust (Senthilkumar et al., 2024; Sarkar, 2024; Celik et al., 2024). Interactional governance also demands uncertainty articulation in an instructionally legible form, so teachers can triage suggestions quickly without performing elaborate verification under time pressure. The co-teacher should operationalize abstention as competence, not as failure, by routing ambiguous or high-stakes content through teacher escalation. These regimes and protocols are enumerated in Table 3 as a pattern library that can be configured across diverse educational levels, disciplines, and resource contexts.

Table 3. Interaction Protocol Patterns for Classroom-Safe Co-Teaching

Interaction Pattern Signature	Trigger Condition and Context Gate	Protocol Sequence and Control Points	Pedagogical Function and Learning Mechanism	Failure Mode and Built-In Safeguard
Suggest-Verify-Deliver Loop	*Teacher requests or approves a momentary support need, low latency window*	*AI proposes, teacher verifies, teacher delivers, AI logs rationale for traceability*	*Reduces planning burden in situ, preserves authority, supports adaptive explanation*	*Automation bias mitigated by mandatory verification gate, uncertainty cues, override visibility*
Probe-Diagnose-Respond Cycle	*Misconception suspicion arises, evidence threshold partially met*	*AI suggests probes, teacher selects, AI aggregates response patterns, teacher chooses re-teaching move*	*Strengthens formative inference, targets conceptual repair, sustains epistemic progress*	*Misdiagnosis mitigated by uncertainty marking, abstain triggers on weak evidence, escalation routing*

Dialogic Talk-Move Orchestration	*Whole-class discussion stalls or becomes monologic, discourse equity risk detected*	*AI generates talk-moves, teacher moderates, AI proposes revoicing options, teacher selects*	*Enhances dialogic reasoning, increases justification norms, supports equitable participation*	*Scripted pseudo-dialogue mitigated by teacher personalization controls, authenticity constraints, latency caps*
Backchannel Noticing and Signal Triage	*Teacher workload saturates, multiple groups require monitoring*	*AI flags signals, teacher triages, AI suggests micro-interventions, teacher acts*	*Expands teacher noticing bandwidth, supports pacing regulation, reduces neglect of quiet learners*	*Surveillance creep mitigated by purpose limitation, minimal data capture, opt-in sensing and logging*
Draft-Edit-Release Feedback Workflow	*Post-activity feedback required, criteria available, time scarcity high*	*AI drafts rubric-grounded feedback, teacher edits, release to learners, AI records edits for learning*	*Improves feedback timeliness and specificity, supports actionable next steps*	*Hallucinated feedback mitigated by evidence anchoring, rubric enforcement, abstain-escalate on missing artifacts*
Abstain-Escalate Safety Protocol	*High-stakes, ambiguous, sensitive, or policy-restricted content emerges*	*AI abstains, explains uncertainty, routes to teacher, logs event for audit*	*Prevents epistemic harm, preserves safeguarding, maintains trust calibration*	*Over-abstention mitigated by calibrated thresholds, teacher-tunable settings, post-event review loop*

Table 3 is intentionally specified as a protocol grammar rather than a feature list, because classroom safety depends on repeatable routines that can be trained, audited, and improved through reflective practice. The patterns also instantiate the theoretical commitments of Table 1 and the authority regimes of Table 2, by embedding verification, uncertainty articulation, and escalation as structural properties of interaction rather than discretionary habits. Importantly, the backchannel pattern is framed as noticing support rather than surveillance, because pedagogical legitimacy requires purpose limitation, data minimization, and teacher control. The next subsections deepen the interaction story by analyzing discourse quality as an epistemic constraint and by formalizing provenance discipline and contestability as core interactional affordances for global, high-accountability schooling environments.

Discourse Engineering, Epistemic Quality, and Participation Equity Orchestration

Co-teaching must be discourse-aware because classroom learning is mediated by language, turn-taking, and normative epistemic moves, not merely by exposure to explanations. A discourse-engineered co-teacher supports epistemic moves such as justification, conjecture, refutation, counterexample generation, and transfer prompting, while preventing premature closure that collapses inquiry into answer consumption (Li et al., 2025; Cukurova, 2025; Marrone et al., 2024). This aligns with dialogic pedagogy and epistemic cognition constructs that treat knowledge as something that must be warranted, not merely stated. Participation equity is an orchestration target, not a moral add-on, because inequitable airtime distribution systematically deprives some learners of cognitive rehearsal and social recognition (Topali et al., 2025; Chan & Tsi, 2024; Luo, 2024). The co-teacher can support equity by surfacing talk-distribution imbalances, proposing inclusive prompts, and suggesting structured turn-taking routines, but it must do so without stigmatizing learners or reducing classroom interaction to behavioral compliance. Table 3's dialogic talk-move protocol provides a bounded channel for this support, ensuring that discourse interventions remain teacher-moderated and context-sensitive. In multilingual settings, discourse engineering must also respect translanguaging and accessibility needs, treating linguistic variability as a resource rather than a deficit.

Provenance Discipline, Uncertainty Articulation and Contestability-by-Design

A co-teacher's instructional legitimacy depends on provenance discipline, because teaching is accountable to curricular standards, assessment criteria, and community norms, and because fluent text without traceability can launder error into authority. Provenance discipline requires that the co-teacher maintain a transparent linkage between suggestions and the curricular intent, task criteria, and approved knowledge sources that justify them, enabling rapid teacher verification under time constraints (Sadeghi & Niu, 2024; Nguyen et al., 2024; Kaswan et al., 2024). Uncertainty articulation must be interactionally legible, differentiating low-confidence suggestions, assumption-heavy inferences, and evidence-missing claims, thereby supporting calibrated reliance and preventing automation bias. Contestability-by-design completes this triad by ensuring that both teachers and learners can challenge AI-mediated guidance through explainable rationale prompts, alternative proposal generation, and escalation pathways that preserve human judgment (Santoso & Wijayanti, 2024; Bulathwela et al., 2024; Francis et al., 2025). Table 3 embeds these constructs through verification gates, abstain-escalate routines, and rubric grounding, but the conceptual point is broader, co-teaching is sustainable only when the interaction layer makes epistemic status visible. This is especially critical in high-stakes assessment contexts and in policy-sensitive topics, where the cost of misalignment is institutional and moral, not merely cognitive.

5. Epistemic Governance, Ethics and Accountability as Instructional Infrastructure

Human-AI co-teaching becomes defensible only when governance is treated as instructional infrastructure, because classrooms are high-accountability micro-institutions in which pedagogical decisions carry legal, ethical, and developmental consequences. Governance must be conceptualized as a set of designable controls over decision rights, data rights, error rights, and appeal rights, rather than as post hoc compliance (Takahashi et al., 2025; Khan, 2024). In a co-teacher ecology, reliability is not an emergent property of model scale, it is a property of sociotechnical constraint, including permissioning, auditability, and bounded autonomy. A governance substrate must therefore specify what the AI is allowed to do, what it must never do, what it must log, and when it must abstain and escalate, consistent with the interaction patterns previously formalized in Table 3 (Yan, L., et al., 2024b; Al-Zahrani & Alasmari, 2024; Butson & Spronken-Smith, 2024). This section operationalizes governance using constructs from value-sensitive design, resilience engineering, and institutional risk management, framing harm as predictable failure modes rather than exceptional incidents. Table 4 will consolidate these risks into a control-primitives register that can guide policy, procurement, and classroom configuration across global contexts.

Epistemic Integrity, Knowledge Legitimacy, and the Hazard of Fluent Wrongness

The central epistemic risk in co-teaching is fluent wrongness, where rhetorically coherent outputs mimic warranted explanation without being anchored to evidence or curricular truth conditions. This is not merely a correctness issue, it is a legitimacy issue because it can restructure learners' epistemic norms, lowering epistemic vigilance and elevating persuasive fluency over justification (Nithithanatchinnapat et al., 2024; Farahani & Ghasemi, 2024; Akpan et al., 2025). Co-teaching must therefore preserve epistemic integrity through provenance discipline, evidence anchoring, and uncertainty articulation, ensuring that explanations and feedback remain tethered to criteria, definitions, and task-specific warrants. Authority conflict is another epistemic hazard, because a co-teacher can inadvertently compete with teacher judgment, curricular commitments, or culturally situated knowledge, producing epistemic dissonance and erosion of trust (Nkechi et al., 2024; Ifenthaler et al., 2024; Rawas, 2024). Epistemic governance must also address representational framing, including overgeneralization, false balance, and normative laundering, particularly in humanities, civics, and social science contexts. These hazards are structurally predictable when verification and contestability are weak, and Table 4 maps them to concrete controls such as abstain-escalate thresholds, rationale visibility, and audit logging.

Equity, Inclusion, and Differential Harm Distribution as First-Order Constraints

Equity is not a secondary evaluation lens in co-teaching, it is a first-order design constraint because error and misalignment are rarely evenly distributed across learners, languages, and accessibility contexts. Differential harm arises when the co-teacher performs better for dominant language varieties, mainstream cultural referents, or normative discourse styles, thereby widening opportunity-to-learn gaps under the guise of personalized support (Bettayeb et al., 2024; Sun et al., 2024; Tariq, 2024). Inclusion also has a structural dimension, because accessibility transformations that simplify language or tasks can unintentionally reduce epistemic demand, producing a hidden tracking effect that undermines dignity and long-term attainment. An equity-governed co-teacher must therefore incorporate fairness constraints across representational choice, discourse participation support, and feedback tone, coupled with contestability pathways that allow learners to challenge mischaracterizations (Yan, L., et al., 2024a; Bansal et al., 2024). Governance must also address stigma risks in grouping and intervention suggestions, because algorithmic grouping rationales can encode deficit framings if not bounded by humane constraints. Table 4 formalizes equity risks as governance objects, linking them to control primitives such as minimal-data configuration, teacher override defaults, and explicit fairness checks in orchestration routines.

Privacy, Data Governance, and the Boundary Between Noticing and Surveillance

Privacy governance is pivotal in co-teaching because instructional systems can convert classroom interaction into persistent data traces, enabling function creep from pedagogical support into surveillance. The distinction between noticing and surveillance is therefore not rhetorical, it is operational, noticing is purpose-limited support for learning regulation, surveillance is expansive monitoring beyond pedagogical necessity (Omran Zailuddin et al., 2024; Wood & Moss, 2024; Kayyali, 2024). Data governance must be anchored in minimization and purpose limitation, specifying the smallest data required for each role contract in Table 2 and each protocol in Table 3, and prohibiting secondary use without explicit authorization. Retention limits, access controls, and deletion rights are not technical details, they are legitimacy safeguards that protect learners, teachers, and institutions from misuse and reputational risk (Marrone et al., 2025; Creely & Blannin, 2025; Nikolopoulou, 2024). Sensitive data regimes are especially salient when learners are minors, when classroom discourse includes personal disclosures, or when multimodal sensing is involved. Privacy governance also interacts with equity, because surveillance tends to fall more heavily on marginalized learners, amplifying disciplinary disparities. Table 4 articulates these privacy pathways as concrete governance risks with enforceable controls.

Table 4. Governance Risks and Control Primitives in Co-Teaching

Governance Risk Domain	Predictable Harm Pathway	Primary Affected Stakeholders	Control Primitive and Enforcement Logic	Accountability Owner Boundary
Epistemic Error and Fluent Wrongness	*Persuasive explanations unmoored from evidence, curricular drift, misconception reinforcement*	*Learners, teachers*	*Provenance discipline, uncertainty articulation, abstain-escalate thresholds, verification gates*	*Vendor capability disclosure, institution configuration, teacher enactment*
Automation Bias and Authority Displacement	*Overreliance, reduced teacher discretion, illegible influence on pedagogical choices*	*Learners, teachers*	*Teacher-in-command defaults, mandatory override visibility, mixed-initiative constraints, audit logging*	*Institution policy, platform defaults, teacher verification routines*
Equity Drift and Differential Service Quality	*Uneven performance across language varieties, disability contexts, cultural referents, hidden tracking*	*Minoritized learners, families*	*Fairness constraints, inclusive discourse supports, contestability pathways, rigor-preservation checks*	*Institution governance, teacher moderation, vendor safety testing*
Privacy Breach and Surveillance Creep	*Excessive data capture, function creep, persistent learner profiling, unauthorized access*	*Learners, teachers, institutions*	*Data minimization, purpose limitation, retention caps, role-based access control, deletion rights*	*Institution data stewardship, vendor security, policy compliance*
Assessment Validity and Integrity Collapse	*Misattributed authorship, feedback laundering, grading distortion, weakened credential meaning*	*Learners, institutions, employers*	*Assessment redesign rules, provenance for feedback, evidence anchoring, restricted modes for high-stakes tasks*	*Institution assessment governance, teacher oversight, platform restrictions*
Safeguarding and Psychological Harm Exposure	*Inappropriate content, manipulative tone, stigmatizing group suggestions, harmful advice*	*Learners, teachers*	*Content safety constraints, sensitive-topic routing, tone governance, escalation to human safeguarding channels*	*Institution safeguarding policy, vendor safety controls, teacher escalation*

Table 4 functions as a governance blueprint that translates abstract ethics into enforceable constraints, thereby making co-teaching compatible with professional accountability and rights-based educational obligations. The register also clarifies that accountability is distributed, vendors must disclose limitations and provide safety controls, institutions must configure and monitor use, and teachers must enact verification and moderation as part of instructional practice. Importantly, the control primitives are designed to be auditable, enabling learning from incidents through post-event review rather than relying on blame-centric reactions. The next subsection extends this governance architecture by tightening accountability allocation and by specifying audit-ready traceability as a prerequisite for legitimate co-teaching at scale.

Accountability Allocation, Audit-Ready Traceability, and Institutional Legitimacy

Accountability allocation in co-teaching must be explicit because ambiguous responsibility produces governance gaps, and governance gaps produce predictable harm. A defensible model partitions accountability across three layers, instructional discretion at the teacher layer, adoption and configuration responsibility at the institution layer and capability-claims and safety-controls responsibility at the vendor layer, with clear interfaces for incident reporting and remediation (Kim, S. J., 2024; Wong & Looi, 2024; Dhillon et al., 2024). Audit-ready traceability is the practical mechanism that makes this partition credible, requiring that co-teacher interventions, overrides, abstentions, and escalation events be recorded in a minimal yet reviewable form that respects privacy constraints. Traceability also supports epistemic integrity, because it allows teachers and institutions to reconstruct why a suggestion was made, what assumptions were used, and how it was resolved (Arar et al., 2024; Giannakos et al., 2025; Kim, J., et al., 2024). Governance must further ensure contestability for learners, enabling appeals against AI-mediated feedback and protecting epistemic dignity through transparent rationales. These accountability mechanisms are not bureaucratic overhead, they are reliability infrastructure that enables institutional learning, continuous improvement, and public trust in credential validity. The design logic already embedded in Table 4 thus becomes the prerequisite for Section 6's implementation maturity model, where governance, pedagogy, and technical configuration are aligned into a scalable adoption pathway.

6. Design Principles and Institutional Implementation Logic

Designing AI as a co-teacher requires a reliability-first architecture in which pedagogical legitimacy is produced by constraint, traceability, and calibrated initiative rather than by generative capacity alone. The foundational principle is teacher-in-command interaction, where suggestion authority is separated from decision and execution authority, preserving professional discretion while still enabling high-velocity support (Duraimutharasan et al., 2025; Maphoto et al., 2024; García-López et al., 2025). A second principle is uncertainty-first communication, in which the system externalizes confidence, assumptions, and evidentiary gaps in instructionally legible forms, enabling rapid verification under classroom time pressure. A third principle is provenance-by-default, where every explanatory move and feedback draft is anchored to task criteria, curricular intent, and approved knowledge sources, thereby preventing fluent wrongness and authority laundering (Lin & Chen, 2025; Tzirides et al., 2024; Lee et al., 2025). A fourth principle is inclusion-first transformation, integrating accessibility and multilingual support as baseline constraints with rigor-preservation checks that prevent hidden tracking. A fifth principle is resilience and repair, operationalizing rollback, abstain-escalate routing, and post-incident learning routines as default behaviors. These principles anticipate implementable adoption pathways, and Table 5 will operationalize them as maturity levels that align configuration, pedagogy, and governance.

Institutional Readiness as Socio-Technical Maturity and Implementation Feasibility

Institutional adoption of co-teaching is not a device procurement problem, it is a sociotechnical maturity problem involving policy coherence, professional learning capacity, assessment legitimacy, and data governance resilience. Readiness can be conceptualized using normalization process theory, where uptake depends on coherence, cognitive participation, collective action, and reflexive monitoring, and using diffusion-of-innovation logic, where adoption follows social legitimacy, compatibility with existing routines, and observable value (Dai et al., 2025; Clegg & Sarkar, 2024). Co-teaching implementations fail when they are treated as plug-and-play tools, because hidden work emerges in verification, exception handling, and integrity management, all of which require organizational scaffolding.

The institution must therefore define approved use-cases, disallowed use-cases, role-based permissions, logging expectations, retention caps, and escalation pathways that integrate safeguarding and academic integrity (Wen et al., 2024; Kim, 2024b; Kim & Cho, 2025). Readiness also requires alignment with curriculum governance, because ungrounded co-teacher advice can drift from standards, and with assessment governance, because credential validity collapses when authorship boundaries are undefined. Table 5 expresses this readiness as staged maturity levels, each with minimum viable governance requirements that can be enacted globally across varied resource constraints.

Professional Learning as Epistemic Hygiene and Orchestration Competence

Teacher professional learning for co-teaching must be framed as epistemic hygiene and orchestration competence, not as prompt-craft, because the central risk is illegible influence rather than insufficient fluency. Epistemic hygiene includes verification routines, triangulation habits, rubric grounding, and disciplined use of provenance, enabling teachers to differentiate plausible text from warranted instruction. Orchestration competence includes timing judgments, discourse facilitation, participation equity regulation, and adaptive differentiation under constraint, ensuring that co-teacher inputs are integrated without destabilizing classroom interaction (Virvou et al., 2024; Nguyen, 2025; Mittal et al., 2024). A competence model should also include ethical discernment, including privacy minimization, equity sensitivity, and safeguarding escalation, because co-teaching expands the teacher's responsibility surface. Professional learning is most sustainable when embedded in collaborative practice, such as lesson study, peer review of AI-mediated materials, and reflective debriefing aligned to audit logs, because these routines transform governance artifacts into learning resources rather than compliance burdens (Pahi et al., 2024; Hwang & Lee, 2025; Jony & Hamim, 2024). A key design constraint is preventing deskilling, which requires that teachers remain authors of pedagogical intent and arbiters of instructional judgment, while the co-teacher operates as a bounded cognitive amplifier under transparent control points.

Curriculum and Assessment Redesign as Preconditions for Legitimate Scaling

Co-teaching cannot be scaled responsibly without curriculum and assessment redesign, because AI alters the epistemic ecology of classrooms and the meaning of evidence of learning. Curriculum integration requires mapping co-teacher roles to learning progressions, ensuring that generated explanations, examples, and tasks preserve conceptual sequencing and do not flatten disciplinary epistemologies (Dhanasekaran, 2025; Xiao et al., 2025). Assessment redesign is essential because unregulated AI assistance collapses construct validity, misattributes authorship, and degrades the signaling function of credentials for workforce development and public trust. A validity-oriented approach treats assessment as a system of evidence, where tasks must elicit reasoning traces that remain attributable to learners, such as in-situ explanation, oral defense, constrained-resource problem solving, and process documentation (Kim, 2024a; Fragiadakis et al., 2024; Luckin, 2025). Feedback governance must also be redesigned, requiring that AI-drafted feedback remain evidence-anchored and criteria-referenced, avoiding stylistic persuasion that can mask weak understanding. Integrity policies should therefore be expressed as assistance boundaries that are pedagogically intelligible, specifying what kinds of AI support are permissible for learning and what kinds invalidate evidence. These constraints align directly with governance primitives in Table 4 and become operational adoption requirements in Table 5.

Table 5. Institutional Maturity Model for Implementing AI Co-Teaching

Adoption Maturity Tier	Technical Configuration and Permissioning	Pedagogical Integration and Workflow Discipline	Governance and Data Stewardship Controls	Accountability and Auditability Artifacts
Tier 1 Exploratory Piloting	*Standalone tools, minimal integration, teacher-controlled access*	*Individual experimentation, low-stakes planning support*	*Basic minimization norms, informal retention awareness*	*Simple usage charter, manual incident notes*
Tier 2 Constrained Classroom Use	*Approved toolset, role-based access, restricted modes for minors*	*Defined low-risk protocols, suggest-verify-deliver routines*	*Purpose limitation, access control, retention caps*	*Logging expectations, escalation routing, override visibility*
Tier 3 Curriculum-Linked Integration	*LMS-linked workflows, grounded corpora, controlled retrieval*	*Curriculum-aligned routines, formative checks, feedback drafting with edits*	*Audit-ready logs, deletion rights, privacy-by-design configuration*	*Role contracts, verification checklists, periodic review reports*
Tier 4 Governed Co-Teaching Operations	*Mixed-initiative constraints, sensing opt-ins, provenance interfaces*	*Orchestration protocols, discourse supports, equity-aware grouping norms*	*Fairness checks, safeguarding pathways, integrity policy alignment*	*Contestability pathways, incident response playbooks, audit sampling*
Tier 5 High-Reliability Co-Teaching Ecosystem	*Full provenance stack, interoperability, policy-enforced guardrails*	*Institution-wide design norms, continuous improvement loops*	*Routine audits, governance dashboards, compliance monitoring*	*Accountability allocation map, post-incident learning briefs, external assurance-ready records*
Tier 6 Audit-Ready Transnational Scaling	*Portable configurations, localization layers, cross-jurisdiction controls*	*Multilingual, accessibility-first co-teaching at scale*	*Harmonized privacy and integrity controls, vendor assurance requirements*	*Cross-site audit comparability, standardized logs, governance certification artifacts*

Table 5 operationalizes co-teaching adoption as staged sociotechnical capability building, where technical configuration is deliberately synchronized with pedagogical workflow discipline and governance robustness. The maturity framing is not a maturity theatre, it is a risk-management logic that recognizes that advanced capabilities without governance increase harm probability and erode institutional legitimacy. Notably, the higher tiers emphasize provenance interfaces and contestability pathways, because scalable co-teaching requires not only operational efficiency but also public defensibility in the face of integrity, privacy, and equity scrutiny. The model is also globally portable because it does not assume uniform infrastructure, it specifies controls that can be implemented at different levels of technical sophistication, including through restricted modes and localized corpora. The next subsection connects this maturity logic to procurement and platform governance, translating institutional requirements into enforceable technical and contractual constraints.

Procurement, Platform Governance, and Sustainability of Co-Teaching Infrastructures

Procurement is a governance instrument, not a purchasing transaction, because vendor choices determine the availability of safety controls, provenance mechanisms, audit logs, and interoperability that make co-teaching defensible. Platform governance must require documented capability limits, configurability of mixed-initiative regimes, restricted modes appropriate for minors, and interfaces for provenance inspection and override logging, otherwise institutions cannot meet accountability obligations (Edwards et al., 2025; Brusilovsky, 2024; Atchley et al., 2024). Interoperability requirements are crucial for sustainability, because co-teaching must integrate with LMS ecosystems, assessment systems, and accessibility tooling without creating vendor lock-in or brittle data silos that undermine auditability and portability. Sustainability also includes operational feasibility, such as latency performance, offline resilience where needed, maintenance burden, and cost predictability, because classroom reliability degrades when systems are intermittently unavailable (Padovano & Cardamone, 2024; Chiu & Rospigliosi, 2025; Nguyen et al., 2024). Institutions must also specify data portability and deletion rights to maintain stewardship across contract changes and jurisdictional constraints. When procurement is aligned to the maturity model in Table 5, platform governance becomes a mechanism for translating pedagogical and ethical commitments into enforceable system properties, thereby enabling co-teaching that is scalable, equitable, and legitimate across diverse global educational contexts.

7. Conclusion

Human-AI collaborative teaching is most coherently understood as hybrid instructional cognition, a joint cognitive system embedded in a sociotechnical institution where legitimacy, reliability, and learning quality are co-produced through constraint, coupling, and accountable orchestration. The conceptual synthesis advanced across Sections 2 through 6 clarifies that AI fluency is neither a sufficient nor a stable basis for co-teaching, because classroom viability depends on interaction architectures, role ecologies, and governance primitives that control epistemic risk, equity drift, privacy exposure, and integrity collapse. Co-teaching becomes a designable phenomenon when it is decomposed into bounded roles with explicit authority allocation, when interaction is specified as protocol rather than improvisation, and when governance is treated as instructional infrastructure that defines decision rights, data rights, error rights, and appeal rights. The integrated architecture also repositions professional agency as a non-delegable normative anchor, requiring teacher-in-command defaults, verification discipline, and abstain-escalate routing as institutionalized competence rather than discretionary caution. In this framing, successful co-teaching is not an adoption event, it is a capability regime that aligns cognition, pedagogy, and governance into a high-reliability instructional stack.

The architecture resolves persistent category confusion by specifying necessary features that distinguish co-teaching from generic tool use, including role explicitness, bounded permissions, provenance discipline, calibrated uncertainty, contestability pathways, and audit-ready traceability. It also resolves the false dichotomy between innovation and accountability by showing that governance controls, such as purpose limitation, retention caps, and override visibility, are enabling conditions for scalable use, not bureaucratic friction. The role ecology model clarifies how division of instructional labour can be reorganized without deskilling, by separating suggestion authority from decision and execution authority and by preserving teacher discretion as the locus of ethical and professional judgment. The interaction pattern library demonstrates how mixed-initiative systems can be stabilized under classroom constraints through protocol grammars that reduce interruption costs, minimize cognitive switching, and support discourse-quality and participation equity without scripting inauthentic dialogue. The maturity model translates these theoretical commitments into institutional implementation logic that is portable across global contexts, including multilingual and low-resource settings, because it specifies governance and workflow requirements as staged capabilities rather than assuming uniform infrastructure. Together, these contributions enable institutions to articulate concrete adoption criteria, procurement constraints, and professional learning targets that are defensible to regulators, families, workforce stakeholders, and academic communities.

A forward-looking co-teaching agenda should be anchored in theoretical commitments that keep the paradigm stable under technological change, including the principle that reliability is a property of coupling and governance, not of model size, and the principle that educational legitimacy is preserved through contestability and traceability, not through opacity and performance theater. Co-teaching should therefore evolve toward richer provenance interfaces, stronger uncertainty articulation, and more refined mixed-initiative regimes that can modulate proactivity by risk class, instructional phase, and learner vulnerability. Equity must remain a first-order constraint through fairness-aware orchestration and rigor-preserving accessibility transformations, because personalization without governance can institutionalize differential opportunity-to-learn and covert tracking. Assessment redesign should be treated as a structural prerequisite, since credential legitimacy is a societal asset and an economic signal, and co-teaching must not erode the validity of evidence of learning. Institutional learning loops should be normalized through post-incident reflection, audit sampling, and continuous professional learning that treats logs as pedagogical artifacts for improvement rather than as surveillance. When these commitments are sustained, the co-teacher paradigm can mature into a globally defensible instructional infrastructure that augments teacher noticing, strengthens formative inference, and expands access to high-quality pedagogy while preserving epistemic dignity, safeguarding, and professional authority.

Declaration of Interest:
No potential conflict of interest was reported by the authors.

Funding Information:
This research did not receive any specific funding from any public, commercial, or non-profit agency.

Disclosure Statement:
No material or relevant stake relating to this research was disclosed by the author(s).

Competing Interest:
No potential conflict of interest was reported by the author(s).

Data Availability Statement:
Data sharing is not applicable to this research article as no new data were created or analysed in this study.

References

- Adel, A., & HS Alani, N. (2024). Human-centric collaboration and industry 5.0 framework in smart cities and communities: fostering sustainable development goals 3, 4, 9, and 11 in society 5.0. *Smart Cities*, *7*(4), 1723-1775.
- Akpan, I. J., Kobara, Y. M., Owolabi, J., Akpan, A. A., & Offodile, O. F. (2025). Conversational and generative artificial intelligence and human–chatbot interaction in education and research. *International Transactions in Operational Research*, *32*(3), 1251-1281.
- Al-Kfairy, M., Mustafa, D., Kshetri, N., Insiew, M., & Alfandi, O. (2024, September). Ethical challenges and solutions of generative AI: An interdisciplinary perspective. In *Informatics* (Vol. 11, No. 3, p. 58). Multidisciplinary Digital Publishing Institute.
- Al-Zahrani, A. M., & Alasmari, T. M. (2024). Exploring the impact of artificial intelligence on higher education: The dynamics of ethical, social, and educational implications. *Humanities and Social Sciences Communications*, *11*(1), 1-12.
- Anane-Simon, R., & Atiku, S. O. (2024). Artificial intelligence and automation for the future of startups. In *Ecosystem Dynamics and Strategies for Startups Scalability* (pp. 133-153). IGI Global Scientific Publishing.
- Arar, K., Tlili, A., & Salha, S. (2024). Human-Machine symbiosis in educational leadership in the era of artificial intelligence (AI): Where are we heading?. *Educational Management Administration & Leadership*, 17411432241292295.
- Asad, M. M., & Ajaz, A. (2024). Impact of ChatGPT and generative AI on lifelong learning and upskilling learners in higher education: unveiling the challenges and opportunities globally. *The International Journal of Information and Learning Technology*, *41*(5), 507-523.
- Atchley, P., Pannell, H., Wofford, K., Hopkins, M., & Atchley, R. A. (2024). Human and AI collaboration in the higher education environment: opportunities and concerns. *Cognitive research: principles and implications*, *9*(1), 20.
- Bansal, G., Mitchell, A., & Li, D. (2024). A Panel Report on Higher Education in the Age of AI from the Perspective of Academic Leaders in the Midwest US. *Communications of the Association for Information Systems*, *54*(1), 360-375.
- Bettayeb, A. M., Abu Talib, M., Sobhe Altayasinah, A. Z., & Dakalbab, F. (2024, July). Exploring the impact of ChatGPT: conversational AI in education. In *Frontiers in Education* (Vol. 9, p. 1379796). Frontiers Media SA.
- Bozkurt, A. (Ed.). (2024). Why generative AI literacy, why now and why it matters in the educational landscape?: Kings, queens and GenAI dragons. *Open Praxis*, *16*(3), 283-290.
- Bozkurt, A., Xiao, J., Farrow, R., Bai, J. Y., Nerantzi, C., Moore, S., ... & Asino, T. I. (Eds.). (2024). The manifesto for teaching and learning in a time of generative AI: A critical collective stance to better navigate the future. *Open Praxis*, *16*(4), 487-513.
- Brusilovsky, P. (2024). AI in education, learner control, and human-AI collaboration. *International Journal of Artificial Intelligence in Education*, *34*(1), 122-135.
- Bulathwela, S., Pérez-Ortiz, M., Holloway, C., Cukurova, M., & Shawe-Taylor, J. (2024). Artificial intelligence alone will not democratise education: On educational inequality, techno-solutionism and inclusive tools. *Sustainability*, *16*(2), 781.
- Butson, R., & Spronken-Smith, R. (2024). AI and its implications for research in higher education: a critical dialogue. *Higher Education Research & Development*, *43*(3), 563-577.
- Cai, L., Msafiri, M. M., & Kangwa, D. (2025). Exploring the impact of integrating AI tools in higher education using the Zone of Proximal Development. *Education and Information Technologies*, *30*(6), 7191-7264.

- Celik, I., Gedrimiene, E., Siklander, S., & Muukkonen, H. (2024). The affordances of artificial intelligence-based tools for supporting 21st-century skills: A systematic review of empirical research in higher education. *Australasian Journal of Educational Technology*, *40*(3), 19-38.
- Chan, C. K. Y., & Tsi, L. H. (2024). Will generative AI replace teachers in higher education? A study of teacher and student perceptions. *Studies in Educational Evaluation*, *83*, 101395.
- Chee, H., Ahn, S., & Lee, J. (2025). A competency framework for AI literacy: Variations by different learner groups and an implied learning pathway. *British Journal of Educational Technology*, *56*(5), 2146-2182.
- Chiu, T. K., & Rospigliosi, P. A. (2025). Encouraging human-AI collaboration in interactive learning environments. *Interactive Learning Environments*, *33*(2), 921-924.
- Chuang, S. (2024). Indispensable skills for human employees in the age of robots and AI. *European Journal of Training and Development*, 48(1/2), 179-195.
- Cingillioglu, I., Gal, U., & Prokhorov, A. (2024). AI-experiments in education: An AI-driven randomized controlled trial for higher education research. *Education and Information Technologies*, *29*(15), 19649-19677.
- Clegg, S., & Sarkar, S. (2024). Artificial intelligence and management education: A conceptualization of human-machine interaction. *The International Journal of Management Education*, *22*(3), 101007.
- Crawford, J., Allen, K. A., Pani, B., & Cowling, M. (2024). When artificial intelligence substitutes humans in higher education: the cost of loneliness, student success, and retention. *Studies in Higher Education*, *49*(5), 883-897.
- Creely, E., & Blannin, J. (2025). Creative partnerships with generative AI. Possibilities for education and beyond. *Thinking Skills and Creativity*, *56*, 101727.
- Cukurova, M. (2025). The interplay of learning, analytics and artificial intelligence in education: A vision for hybrid intelligence. *British Journal of Educational Technology*, *56*(2), 469-488.
- Dai, R., Thomas, M. K. E., & Rawolle, S. (2025). The roles of AI and educational leaders in AI-assisted administrative decision-making: a proposed framework for symbiotic collaboration. *The Australian Educational Researcher*, *52*(2), 1471-1487.
- Dhanasekaran, R. (2025). AI Revolutionizing Education: Technology and Human Interaction in Partnership. *Journal of Computer Science and Technology Studies*, *7*(6), 501-506.
- Dhillon, P. S., Molaei, S., Li, J., Golub, M., Zheng, S., & Robert, L. P. (2024, May). Shaping human-AI collaboration: Varied scaffolding levels in co-writing with language models. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (pp. 1-18).
- Duraimutharasan, N., Deepan, A., Swadhi, R., Velmurugan, P. R., & Varshney, K. R. (2025). Enhancing Control Engineering Through Human-Machine Collaboration: AI for Improved Efficiency and Decision-Making. In *Harnessing AI for Control Engineering* (pp. 155-176). IGI Global Scientific Publishing.
- Edwards, J., Nguyen, A., Lämsä, J., Sobocinski, M., Whitehead, R., Dang, B., ... & Järvelä, S. (2025). Human-AI collaboration: Designing artificial agents to facilitate socially shared regulation among learners. *British Journal of Educational Technology*, *56*(2), 712-733.
- Fan, Y., Tang, L., Le, H., Shen, K., Tan, S., Zhao, Y., ... & Gašević, D. (2025). Beware of metacognitive laziness: Effects of generative artificial intelligence on learning motivation, processes, and performance. *British Journal of Educational Technology*, *56*(2), 489-530.
- Farahani, M., & Ghasemi, G. (2024). Artificial intelligence and inequality: Challenges and opportunities. *Int. J. Innov. Educ*, *9*, 78-99.
- Fissore, C., Floris, F., Conte, M. M., & Sacchet, M. (2024, March). Teacher training on artificial intelligence in education. In *Smart Learning Environments in the Post Pandemic Era: Selected Papers from the CELDA 2022 Conference* (pp. 227-244). Cham: Springer Nature Switzerland.
- Fragiadakis, G., Diou, C., Kousiouris, G., & Nikolaidou, M. (2024). Evaluating human-ai collaboration: A review and methodological framework. *arXiv preprint arXiv:2407.19098*.
- Francis, N. J., Jones, S., & Smith, D. P. (2025). Generative AI in higher education: Balancing innovation and integrity. *British Journal of Biomedical Science*, *81*, 14048.
- Fu, Y., & Weng, Z. (2024). Navigating the ethical terrain of AI in education: A systematic review on framing responsible human-centered AI practices. *Computers and Education: Artificial Intelligence*, *7*, 100306.
- García-López, I. M., González, C. S. G., Ramírez-Montoya, M. S., & Molina-Espinosa, J. M. (2025). Challenges of implementing ChatGPT on education: Systematic literature review. *International Journal of Educational Research Open*, *8*, 100401.
- Giannakos, M., Azevedo, R., Brusilovsky, P., Cukurova, M., Dimitriadis, Y., Hernandez-Leo, D., ... & Rienties, B. (2025). The promise and challenges of generative AI in education. *Behaviour & Information Technology*, *44*(11), 2518-2544.
- Goldman, S. R., Taylor, J., Carreon, A., & Smith, S. J. (2024). Using AI to support special education teacher workload. *Journal of Special Education Technology*, *39*(3), 434-447.
- Groothuijsen, S., van den Beemt, A., Remmers, J. C., & van Meeuwen, L. W. (2024). AI chatbots in programming education: Students' use in a scientific computing course and consequences for learning. *Computers and Education: Artificial Intelligence*, *7*, 100290.

- Guggemos, J. (2024). To fear or not to fear–Human resource development professionals' positioning towards artificial intelligence with a focus on augmentation. *Computers and Education: Artificial Intelligence*, *7*, 100260.
- Gupta, N. (2025). The Rise of AI Copilots: Redefining Human-Machine Collaboration in Knowledge Work. *International Journal of Humanities and Information Technology*, *7*(03).
- Han, A., Zhou, X., Cai, Z., Han, S., Ko, R., Corrigan, S., & Peppler, K. A. (2024, May). Teachers, parents, and students' perspectives on integrating generative AI into elementary literacy education. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (pp. 1-17).
- Hutson, J. (2024). Rethinking plagiarism in the era of generative AI. *Journal of Intelligent Communication*, *4*(1).
- Hwang, Y., & Lee, J. H. (2025). Exploring students' experiences and perceptions of human-AI collaboration in digital content making. *International Journal of Educational Technology in Higher Education*, 22(1), 44.
- Ifenthaler, D., Majumdar, R., Gorissen, P., Judge, M., Mishra, S., Raffaghelli, J., & Shimada, A. (2024). Artificial intelligence in education: Implications for policymakers, researchers, and practitioners. *Technology, Knowledge and Learning*, *29*(4), 1693-1710.
- Imran, M., Almusharraf, N., Abdellatif, M. S., & Abbasova, M. Y. (2024). Artificial intelligence in higher education: Enhancing learning systems and transforming educational paradigms.
- Jacques, P. H., Moss, H. K., & Garger, J. (2024). A synthesis of AI in higher education: Shaping the future. *Journal of Behavioral and Applied Management*, *24*(2), 103-111.
- Jiang, T., Sun, Z., Fu, S., & Lv, Y. (2024). Human-AI interaction research agenda: A user-centered perspective. *Data and Information Management*, *8*(4), 100078.
- Jony, A. I., & Hamim, S. A. (2024). Empowering virtual collaboration: harnessing AI for enhanced teamwork in higher education. *Educational Technology Quarterly*, *2024*(3), 337-359.
- Joseph, O. B., & Uzondu, N. C. (2024). Integrating AI and Machine Learning in STEM education: Challenges and opportunities. *Computer Science & IT Research Journal*, *5*(8), 1732-1750.
- Jung, D., & Suh, S. (2024). Enhancing soft skills through generative AI in sustainable fashion textile design education. *Sustainability*, *16*(16), 6973.
- Karakose, T., & Tülübas, T. (2024). School Leadership and Management in the Age of Artificial Intelligence (AI): Recent Developments and Future Prospects. *Educational Process: International Journal*, *13*(1), 7-14.
- Kaspersen, M. H., Musaeus, L. H., Bilstrup, K. E. K., Petersen, M. G., Iversen, O. S., Dindler, C., & Dalsgaard, P. (2024, May). From primary education to premium workforce: Drawing on K-12 approaches for developing AI literacy. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (pp. 1-16).
- Kaswan, K. S., Dhatterwal, J. S., & Ojha, R. P. (2024). AI in personalized learning. In *Advances in technological innovations in higher education* (pp. 103-117). CRC Press.
- Kayyali, M. (2024). Future possibilities and challenges of AI in education. In *Transforming education with generative AI: Prompt engineering and synthetic content creation* (pp. 118-137). IGI Global Scientific Publishing.
- Khan, W. N. (2024). Ethical challenges of AI in education: Balancing innovation with data privacy. *AI EDIFY Journal*, *1*(1), 1-13.
- Kim, J. (2024). Leading teachers' perspective on teacher-AI collaboration in education. *Education and information technologies*, 29(7), 8693-8724.
- Kim, J. (2024). Types of teacher-AI collaboration in K-12 classroom instruction: Chinese teachers' perspective. *Education and Information Technologies*, *29*(13), 17433-17465.
- Kim, J., & Cho, Y. H. (2025). My teammate is AI: Understanding students' perceptions of student-AI collaboration in drawing tasks. *Asia Pacific Journal of Education*, *45*(3), 1013-1027.
- Kim, J., Ham, Y., & Lee, S. S. (2024). Differences in student-AI interaction process on a drawing task: Focusing on students' attitude towards AI and the level of drawing skills. *Australasian Journal of Educational Technology*, *40*(1), 19-41.
- Kim, J., Yu, S., Detrick, R., & Li, N. (2025). Exploring students' perspectives on Generative AI-assisted academic writing. *Education and Information Technologies*, *30*(1), 1265-1300.
- Kim, S. J. (2024). Generative artificial intelligence in collaborative ideation: Educational insight from fashion students. *IEEE Access*, *12*, 49261-49274.
- Kim, T., Shin, D., Kim, Y. H., & Hong, H. (2024, May). DiaryMate: Understanding user perceptions and experience in human-AI collaboration for personal journaling. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (pp. 1-15).
- Kolbjørnsrud, V. (2024). Designing the Intelligent Organization: six principles for Human-AI collaboration. *California Management Review*, *66*(2), 44-64.
- Laak, K. J., & Aru, J. (2025). AI and personalized learning. *Educational Technology & Society*, *28*(4), 133-150.
- Lata, P. (2024). Beyond algorithms: Humanizing artificial intelligence for personalized and adaptive learning. *International Journal of Innovative Research in Engineering and Management*, *11*(5), 10-55524.
- Lee, G. G., Mun, S., Shin, M. K., & Zhai, X. (2025). Collaborative learning with artificial intelligence speakers: pre-service elementary science teachers' responses to the prototype. *Science & Education*, *34*(2), 847-875.
- Leong, W. Y., & Zhang, J. B. (2025). Ethical design of AI for education and learning systems. *ASM Science Journal*, *20*(1), 1-9.

- Li, H., Wang, Y., & Qu, H. (2024, May). Where are we so far? understanding data storytelling tools from the perspective of human-ai collaboration. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (pp. 1-19).
- Li, Y., Zhou, X., Yin, H. B., & Chiu, T. K. (2025). Design language learning with artificial intelligence (AI) chatbots based on activity theory from a systematic review. *Smart Learning Environments*, *12*(1), 24.
- Li, Z., Dhruv, A., & Jain, V. (2024, February). Ethical considerations in the use of AI for higher education: A comprehensive guide. In *2024 IEEE 18th International Conference on Semantic Computing (ICSC)* (pp. 218-223). IEEE.
- Liang, E. S., & Bai, S. (2025). Generative AI and the future of connectivist learning in higher education. *Journal of Asian Public Policy*, *18*(2), 329-351.
- Lin, C. P., & Chen, P. C. (2025). Mentoring for effective human-AI collaboration: an integrated theoretical framework. *Total Quality Management & Business Excellence*, 1-15.
- Łodzikowski, K., Foltz, P. W., & Behrens, J. T. (2024). Generative AI and its educational implications. In *Trust and Inclusion in AI-Mediated Education: Where Human Learning Meets Learning Machines* (pp. 35-57). Cham: Springer Nature Switzerland.
- Luckin, R. (2025). Nurturing human intelligence in the age of AI: rethinking education for the future. *Development and Learning in Organizations: An International Journal*, *39*(1), 1-4.
- Luo, J. (2024). A critical review of GenAI policies in higher education assessment: A call to reconsider the "originality" of students' work. *Assessment & Evaluation in Higher Education*, *49*(5), 651-664.
- Mao, J., Chen, B., & Liu, J. C. (2024). Generative artificial intelligence in education and its implications for assessment. *TechTrends*, *68*(1), 58-66.
- Maphoto, K. B., Sevnarayan, K., Mohale, N. E., Suliman, Z., Ntsopi, T. J., & Mokoena, D. (2024). Advancing students' academic excellence in distance education: Exploring the potential of generative AI integration to improve academic writing skills. *Open Praxis*, *16*(2), 142-159.
- Marrone, R., Cropley, D., & Medeiros, K. (2024). How does narrow AI impact human creativity?. *Creativity Research Journal*, 1-11.
- Marrone, R., Zamecnik, A., Joksimovic, S., Johnson, J., & De Laat, M. (2025). Understanding student perceptions of artificial intelligence as a teammate. *Technology, Knowledge and Learning*, *30*(3), 1847-1869.
- Melweth, H. M. A., Alkahtani, A. S., & Al Mdawi, A. M. M. (2024). The impact of artificial intelligence on improving the quality of education and reducing future anxiety among a sample of teachers in Saudi Arabia. *Kurdish Studies*, *12*(2), 5741-5758.
- Milana, M., Brandi, U., Hodge, S., & Hoggan-Kloubert, T. (2024). Artificial intelligence (AI), conversational agents, and generative AI: Implications for adult education practice and research. *International Journal of Lifelong Education*, *43*(1), 1-7.
- Mittal, U., Sai, S., Chamola, V., & Sangwan, D. (2024). A comprehensive review on generative AI for education. *IEEE Access*, *12*, 142733-142759.
- Mustafa, M. Y., Tlili, A., Lampropoulos, G., Huang, R., Jandrić, P., Zhao, J., ... & Saqr, M. (2024). A systematic review of literature reviews on artificial intelligence in education (AIED): a roadmap to a future research agenda. *Smart Learning Environments*, *11*(1), 59.
- Nguyen, A., Hong, Y., Dang, B., & Huang, X. (2024). Human-AI collaboration patterns in AI-assisted academic writing. *Studies in Higher Education*, *49*(5), 847-864.
- Nguyen, A., Kremantzis, M., Essien, A., Petrounias, I., & Hosseini, S. (2024). Enhancing student engagement through artificial intelligence (AI): Understanding the basics, opportunities, and challenges. *Journal of University Teaching and Learning Practice*, *21*(6), 1-13.
- Nguyen, K. V. (2025). The use of generative AI tools in higher education: Ethical and pedagogical principles. *Journal of Academic Ethics*, 1-21.
- Nikolopoulou, K. (2024). Generative artificial intelligence in higher education: Exploring ways of harnessing pedagogical practices with the assistance of ChatGPT. *International Journal of Changes in Education*, *1*(2), 103-111.
- Nithithanatchinnapat, B., Maurer, J., Deng, X., & Joshi, K. D. (2024). Future business workforce: Crafting a generative AI-centric curriculum today for tomorrow's business education. *ACM SIGMIS Database: The DATABASE for Advances in Information Systems*, *55*(1), 6-11.
- Nkechi, A. A., Ojo, A. O., & Eneh, O. A. (2024). Impact of artificial intelligence in achieving quality education. *Artificial Intelligence and Education-Shaping the Future of Learning: Shaping the Future of Learning*, *73*.
- Noroozi, O., Soleimani, S., Farrokhnia, M., & Banihashem, S. K. (2024). Generative AI in Education: Pedagogical, Theoretical, and Methodological Perspectives. *International Journal of Technology in Education*, *7*(3), 373-385.
- Nurhasanah, S., & Nugraha, M. S. (2024). The future of learning: ethical and philosophical implications of artificial intelligence (AI) integration in education. *Indonesian Journal of Multidiciplinary Research*, *4*(2), 341-352.
- Omran Zailuddin, M. F. N., Nik Harun, N. A., Abdul Rahim, H. A., Kamaruzaman, A. F., Berahim, M. H., Harun, M. H., & Ibrahim, Y. (2024). Redefining creative education: a case study analysis of AI in design courses. *Journal of Research in Innovative Teaching & Learning*, *17*(2), 282-296.
- Padovano, A., & Cardamone, M. (2024). Towards human-AI collaboration in the competency-based curriculum development process: The case of industrial engineering and management education. *Computers and Education: Artificial Intelligence*, *7*, 100256.
- Pahi, K., Hawlader, S., Hicks, E., Zaman, A., & Phan, V. (2024). Enhancing active learning through collaboration between human teachers and generative AI. *Computers and Education Open*, *6*, 100183.

- Park, H., & Ahn, D. (2024, May). The promise and peril of ChatGPT in higher education: opportunities, challenges, and design implications. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (pp. 1-21).
- Peláez-Sánchez, I. C., Velarde-Camaqui, D., & Glasserman-Morales, L. D. (2024, June). The impact of large language models on higher education: exploring the connection between AI and Education 4.0. In *Frontiers in Education* (Vol. 9, p. 1392091). Frontiers Media SA.
- Peters, M. A., & Green, B. J. (2024). Wisdom in the Age of AI Education. *Postdigital science and education*, *6*(4), 1173-1195.
- Pratschke, B. M. (2024). *Generative AI and education: Digital pedagogies, teaching innovation and learning design*. Springer Nature.
- Przegalinska, A., & Triantoro, T. (2024). *Converging minds: The creative potential of collaborative AI*. CRC Press.
- Qian, C., & Wexler, J. (2024, March). Take it, leave it, or fix it: Measuring productivity and trust in human-ai collaboration. In *Proceedings of the 29th International Conference on Intelligent User Interfaces* (pp. 370-384).
- Rane, N., & Choudhary, S. (2024). Role and challenges of ChatGPT, Google Bard, and similar generative Artificial Intelligence in Arts and Humanities. *Studies in Humanities and Education*, *5*(1), 1-11.
- Rawas, S. (2024). AI: the future of humanity. *Discover artificial intelligence*, *4*(1), 25.
- Ruiz-Rojas, L. I., Salvador-Ullauri, L., & Acosta-Vargas, P. (2024). Collaborative working and critical thinking: Adoption of generative artificial intelligence tools in higher education. *Sustainability*, *16*(13), 5367.
- Sadeghi, S., & Niu, C. (2024). Augmenting human decision-making in K-12 education: the role of artificial intelligence in assisting the recruitment and retention of teachers of color for enhanced diversity and inclusivity. *Leadership and Policy in Schools*, 1-21.
- Sajja, R., Ramirez, C. E., Li, Z., Demiray, B. Z., Sermet, Y., & Demir, I. (2024). Integrating generative AI in hackathons: Opportunities, challenges, and educational implications. *Big Data and Cognitive Computing*, *8*(12), 188.
- Santoso, B., & Wijayanti, R. (2024). Human-AI Collaboration in Creative Industries: Workflows in Media Production and Community-Driven Platforms. *Transactions on Artificial Intelligence, Machine Learning, and Cognitive Systems*, *9*(11), 11-26.
- Sarkar, A. (2024). AI Should Challenge, Not Obey. *Communications of the ACM*, *67*(10), 18-21.
- Senthilkumar, K. R., Jagajeevan, R., & Sangeetha, S. (2024). Impact of AI on library and information science in higher institutions in India: A comprehensive analysis of technological integration and educational implications. In *AI-assisted library reconstruction* (pp. 21-33). IGI Global Scientific Publishing.
- Sowa, K., & Przegalinska, A. (2025). From expert systems to generative artificial experts: A new concept for human-ai collaboration in knowledge work. *Journal of Artificial Intelligence Research*, *82*, 2101-2124.
- Storozhyk, M. (2024). Philosophy of future: Analytical overview of interaction between education, science, and artificial intelligence in the context of contemporary challenges. *Futurity Philosophy*, *3*(1), 23-47.
- Sun, C., Major, L., Daltry, R., Moustafa, N., & Friedberg, A. (2024, July). Teacher-AI collaboration in content recommendation for digital personalised learning among pre-primary learners in Kenya. In *Proceedings of the Eleventh ACM Conference on Learning@ Scale* (pp. 346-350).
- Suryanarayana, K. S., Kandi, V. P., Pavani, G., Rao, A. S., Rout, S., & Krishna, T. S. R. (2024). Artificial intelligence enhanced digital learning for the sustainability of education management system. *The Journal of High Technology Management Research*, *35*(2), 100495.
- Takahashi, N., Onuorah, A., Reznitskaya, A., Chukharev, E., Sykes, A., Flammia, M., & Oyler, J. (2025, October). Evaluating the Reliability of Human–AI Collaborative Scoring of Written Arguments Using Rational Force Model. In *Proceedings of the Artificial Intelligence in Measurement and Education Conference (AIME-Con): Works in Progress* (pp. 135-140).
- Tariq, M. U. (2024). Generative AI in curriculum development in higher education. In *Impacts of generative AI on creativity in higher education* (pp. 227-258). IGI Global.
- Tariq, M. U. (2025). AI-Enhanced Project-Based Learning: Revolutionizing Commerce Education With Generative AI. In *Prompt Engineering and Generative AI Applications for Teaching and Learning* (pp. 125-142). IGI Global Scientific Publishing.
- Tarisayi, K. S. (2024, March). Strategic leadership for responsible artificial intelligence adoption in higher education. In *CTE workshop proceedings* (Vol. 11, pp. 4-14).
- Topali, P., Ortega-Arranz, A., Rodríguez-Triana, M. J., Er, E., Khalil, M., & Akçapınar, G. (2025). Designing human-centered learning analytics and artificial intelligence in education solutions: a systematic literature review. *Behaviour & Information Technology*, *44*(5), 1071-1098.
- Tzirides, A. O. O., Zapata, G., Kastania, N. P., Saini, A. K., Castro, V., Ismael, S. A., ... & Kalantzis, M. (2024). Combining human and artificial intelligence for enhanced AI literacy in higher education. *Computers and Education Open*, *6*, 100184.
- Virvou, M., Tsihrintzis, G. A., & Tsichrintzi, E. A. (2024). VIRTSI: A novel trust dynamics model enhancing Artificial Intelligence collaboration with human users–Insights from a ChatGPT evaluation study. *Information Sciences*, *675*, 120759.
- Wang, L., Song, R., Guo, W., & Yang, H. (2025). Exploring prompt pattern for generative artificial intelligence in automatic question generation. *Interactive Learning Environments*, *33*(3), 2559-2584.
- Wen, Q., Liang, J., Sierra, C., Luckin, R., Tong, R., Liu, Z., ... & Tang, J. (2024, August). AI for education (AI4EDU): Advancing personalized education with LLM and adaptive learning. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining* (pp. 6743-6744).

- Wong, L. H., & Looi, C. K. (2024). Advancing the generative AI in education research agenda: Insights from the Asia-Pacific region. *Asia Pacific Journal of Education*, *44*(1), 1-7.
- Wood, D., & Moss, S. H. (2024). Evaluating the impact of students' generative AI use in educational contexts. *Journal of Research in Innovative Teaching & Learning*, *17*(2), 152-167.
- Wu, D., Zhang, S., Ma, Z., Yue, X. G., & Dong, R. K. (2024). Unlocking potential: Key factors shaping undergraduate self-directed learning in AI-enhanced educational environments. *Systems*, *12*(9), 332.
- Xiao, C., Ma, W., Song, Q., Xu, S. X., Zhang, K., Wang, Y., & Fu, Q. (2025, March). Human-ai collaborative essay scoring: A dual-process framework with llms. In *Proceedings of the 15th International Learning Analytics and Knowledge Conference* (pp. 293-305).
- Yadav, S. (2025). Leveraging AI to enhance teaching and learning in education: The role of artificial intelligence in modernizing classroom practices. In *Optimizing research techniques and learning strategies with digital technologies* (pp. 211-238). IGI Global Scientific Publishing.
- Yadav, U., & Shrawankar, U. (2025). Artificial intelligence across industries: A comprehensive review with a focus on education. *AI applications and strategies in teacher education*, 275-320.
- Yan, L., Greiff, S., Teuber, Z., & Gašević, D. (2024). Promises and challenges of generative artificial intelligence for human learning. *Nature Human Behaviour*, *8*(10), 1839-1850.
- Yan, L., Martinez-Maldonado, R., & Gasevic, D. (2024, March). Generative artificial intelligence in learning analytics: Contextualising opportunities and challenges through the learning analytics cycle. In *Proceedings of the 14th learning analytics and knowledge conference* (pp. 101-111).
- Yan, W., Nakajima, T., & Sawada, R. (2024). Benefits and challenges of collaboration between students and conversational generative artificial intelligence in programming learning: An empirical case study. *Education Sciences*, *14*(4), 433.
- Yeter, I. H., Yang, W., & Sturgess, J. B. (2024). Global initiatives and challenges in integrating artificial intelligence literacy in elementary education: Mapping policies and empirical literature. *Future in Educational Research*, *2*(4), 382-402.
- Yim, I. H. Y., & Su, J. (2025). Artificial intelligence (AI) learning tools in K-12 education: A scoping review. *Journal of Computers in Education*, *12*(1), 93-131.
- Yu, J. H., Chauhan, D., Iqbal, R. A., & Yeoh, E. (2025). Mapping academic perspectives on AI in education: Trends, challenges, and sentiments in educational research (2018–2024). *Educational technology research and development*, *73*(1), 199-227.
- Zeb, A., Rehman, F. U., Bin Othayman, M., & Rabnawaz, M. (2025). Artificial intelligence and ChatGPT are fostering knowledge sharing, ethics, academia and libraries. *The International Journal of Information and Learning Technology*, *42*(1), 67-83.