Artificial Intelligence Education Studies
Volume 1, Issue 3, 2025, Page 14–27
https://doi.org/10.6914/aiese.010302

Artificial Intelligence Education Studies

Designing Human-AI Orchestrated Classrooms: Mechanisms, Protocols, and Governance for Competency-Based Education

Xin HUANG
Hong Kong Integration Research Institute

Abstract
The pedagogical promise of Competency-Based Education (CBE) has been historically undermined by profound challenges of scalability, creating an implementation gap between its theoretical merits and practical application. This paper proposes a testable mechanism model wherein Artificial Intelligence (AI) enables the scaling of CBE through three interconnected pathways—diagnostic tracking, adaptive supply, and teacher orchestration—formalized within a distributed cognition framework. To operationalize this model, this paper introduces novel constructs including the "Adaptive-Autonomy Curve" for systematically cultivating self-regulated learning in personalized environments, and a "Situated Performance-Based Assessment Pipeline" for authentic, scalable evaluation of complex skills. The primary contributions of this work are fourfold: first, it provides a rigorous conceptual taxonomy that delineates CBE from adjacent paradigms such as mastery learning and personalized learning; second, it advances a set of falsifiable propositions to guide future empirical research; third, it formalizes the human-AI pedagogical relationship with operational design principles; and fourth, it presents an integrated governance and interoperability protocol for the responsible and effective implementation of AI in competency-based systems.

Keywords competency-based education (CBE); human AI orchestration; adaptive-autonomy curve (self-regulated learning);situated performance-based assessment; governance & interoperability protocols (shared student model)

1 Introduction

The rapid proliferation and increasing sophistication of Artificial Intelligence (AI), particularly generative models, have brought global education systems to a critical inflection point.[1] Initial discourse, often reactive, centered on mitigating threats to academic integrity.[1] However, a more forward-looking perspective recognizes AI not merely as a set of tools to be cautiously adopted, but as a catalyst for the fundamental re-architecting of pedagogy, assessment, and institutional design. This moment demands a strategic shift from reactive policy to proactive, visionary frameworks that harness AI's capabilities to address long-standing, intractable challenges in education.[1]

1.1 The Central Problem: The Scalability Paradox of CBE

For decades, educational reformers have championed a transition from the industrial-era, content-transmission model of schooling to a learner-centric, competency-cultivation paradigm.[1] Competency-Based Education (CBE) represents a leading articulation of this ideal, prioritizing the demonstrated mastery of explicit skills and knowledge over the accumulation of "seat time".[2] The pedagogical benefits are well-documented: CBE ensures foundational knowledge is secure before learners advance, accommodates diverse learning paces, and emphasizes the application of knowledge in authentic contexts, thereby closing learning gaps and enhancing relevance.[4] Despite these advantages, widespread implementation of CBE has been historically constrained by what can be termed the "scalability

© The Authors 2025. Article History Received: June 6, 2025; Accpted: July 19, 2025; Published: September 30, 2025. ***Artificial Intelligence Education Studies***, ISSN 3079-8086 (print), ISSN 3079-8094 (online), a Quarterly, founded on 2025, published by Creative Publishing Co., Limited. Email: wtocom@gmail.com, https://ai-es.org, https://cpcl.hk.

paradox": the very elements that make CBE effective—deep personalization, continuous formative assessment, flexible pacing, and differentiated support—impose an exponential cognitive and logistical load on educators.[7] In a traditional classroom, it is practically impossible for a single teacher to simultaneously track the mastery of dozens of competencies for each of thirty or more students, diagnose individual misconceptions in real time, and provide tailored instructional resources accordingly. This has relegated true CBE to the margins of mainstream education, an aspirational goal perpetually hindered by practical barriers to implementation.[9]

1.2 Thesis Statement

This paper posits that AI can resolve the scalability paradox not as a monolithic 'tool,' but as an orchestrating agent within a re-architected educational system. We propose a testable mid-range theory detailing the specific mechanisms through which AI enables the core functions of CBE at scale, and we formalize the pedagogical, assessment, and governance structures required to support this transformation. This transition is not one of technological determinism but of intentional, systemic design. When integrated into a flawed system, AI risks amplifying existing inequities; however, when implemented as part of a deliberate redesign, it can serve as a powerful catalyst for a more effective and equitable educational paradigm.[1]

1.3 Research Gaps and Contributions

This paper addresses critical gaps in the existing literature and makes four primary contributions to the field of AI in education. First, regarding conceptual clarification, it rectifies the prevalent conceptual ambiguity in the literature by providing a rigorous taxonomy that delineates CBE from mastery learning, standards-based grading, and personalized learning, offering a reusable definitional framework for researchers and practitioners. Second, for mechanistic modeling, it moves beyond descriptive claims by proposing a formal, testable model of how AI scales CBE through three interconnected pathways, yielding a set of falsifiable propositions to guide future empirical research. Third, in operationalizing pedagogy, it formalizes the human-AI pedagogical relationship as a distributed cognitive system and introduces the "Adaptive-Autonomy Curve," a novel design principle to resolve the inherent tension between AI-driven personalization and the development of learner self-regulation. Finally, by bridging assessment and governance, it proposes a concrete "Situated Performance-Based Assessment Pipeline" and a modular governance framework that maps implementation practices to established international standards (e.g., NIST AI RMF, ISO 42001), bridging the gap between innovative pedagogy and the practical demands of ethical, scalable deployment.

2 Conceptual Foundations: A Taxonomy for Learner-Centric Paradigms

To build a rigorous theoretical model, it is first necessary to establish clear and distinct definitions for the foundational concepts that are often used interchangeably in educational discourse. This section addresses this conceptual ambiguity by providing authoritative, operational definitions for Competency-Based Education and its related paradigms.

2.1 Defining Competency-Based Education (CBE)

Drawing from guidance by the U.S. Department of Education and leading consortia, CBE is best understood as a systemic educational framework. The U.S. Department of Education defines it as an approach that "organizes academic content according to competencies—what a student knows and can do—rather than following a more traditional scheme, such as by course".[4] The central tenet is that student progression is based on demonstrated mastery of explicit, measurable learning objectives, not on time spent in a classroom.[2] This model makes time the variable and learning the constant, inverting the traditional structure of education.[5] For the purposes of this paper, the following operational definition is adopted: CBE is a systemic educational framework where the primary unit of academic progression is the verified demonstration of a predefined competency. In this model, time is a variable, while mastery is the constant.

2.2 Delineating Related Constructs

CBE is the overarching system, but its implementation relies on specific strategies, methods, and approaches.

Mastery Learning, first formally proposed by Benjamin Bloom, is an instructional strategy.[11] Its core principle is that students must achieve a high level of proficiency (e.g., 90% accuracy) on a prerequisite unit of knowledge before advancing to new material.[12] The process involves a cycle of initial group instruction, followed by a formative assessment. Students who demonstrate mastery proceed to enrichment activities, while those who do not receive targeted "corrective" instruction and are reassessed until mastery is achieved.[14] Mastery learning is thus a specific pedagogical process for ensuring proficiency within a learning sequence.

Standards-Based Grading (SBG) is an assessment and reporting system designed to communicate student achievement with greater clarity and accuracy.[16] Unlike traditional grading, which often conflates academic performance with factors like effort, participation, or homework completion into a single omnibus grade, SBG reports student performance directly against specific, predefined learning standards or competencies.[16] Grades are based on a body of evidence demonstrating proficiency on a standard, not on an average of points from various assignments.[16] It is the method by which mastery is documented and communicated.

Personalized Learning is a broad instructional approach wherein "the pace of learning and the instructional approach are optimized for the needs of each learner".[18] It involves tailoring learning objectives, content, method, and pace to the individual student's needs, interests, aspirations, and cultural background.[20] While often enabled by technology, personalization is the overarching philosophy of tailoring the educational experience to the individual learner.[20]

The term *"competency"* itself requires precise definition. Moving beyond a simple synonym for "skill," this paper adopts the robust definition from the Organisation for Economic Co-operation and Development (OECD). A competency is "more than just knowledge and skills. It involves the ability to meet complex demands, by drawing on and mobilising psychosocial resources (including skills and attitudes) in a particular context".[23] Similarly, UNESCO defines competency frameworks as structured inventories of expected behaviors, skills, and attitudes that lead to successful performance.[24] A competency, therefore, is not the isolated ability to perform a task but the integrated capacity to apply knowledge and skills effectively in complex, real-world situations.

These definitions reveal a nested, hierarchical relationship. CBE is the overarching systemic framework. Within a CBE system, mastery learning can be the dominant instructional process used to ensure proficiency. The evidence of that mastery is then documented and communicated through a standards-based grading and reporting method. Finally, the delivery of instruction to achieve mastery at an individual level is executed through a personalized learning approach. This clarification is essential, as it allows for a more nuanced analysis of AI's role. AI does not simply "enable CBE"; rather, it provides the technological substrate to automate the mastery learning loop, manage the data infrastructure for SBG, and execute personalization at a scale previously unattainable, all in service of the systemic goals of CBE. To provide a clear reference for these distinctions, the following table summarizes the core attributes of each paradigm.

3 A Testable Mechanism Model: Three Pathways to Scaling CBE with AI

To move beyond the descriptive claim that AI is an "engine for CBE," this section proposes a testable, mid-range theoretical model. This model articulates three distinct but interconnected pathways through which AI systems can resolve the "scalability paradox" and enable the core functions of CBE. These pathways are not independent; they are coupled by a central data architecture—the Shared Student Model—which serves as the dynamic, canonical record of a learner's evolving competencies. This model transforms the abstract potential of AI into a set of concrete, falsifiable mechanisms.

3.1 Diagnostic and Tracking Pathway: The "See" Function

This pathway addresses the fundamental challenge of knowing where each learner is in their journey toward mastery. In a traditional classroom, this function is a primary source of cognitive overload for the teacher.

The mechanism for this pathway involves AI-powered systems continuously collecting and analyzing fine-grained trace data from every learner interaction—every mouse click, keystroke, time-on-task, error pattern, and help-seeking behavior within a digital learning environment.[1] This stream of data, which constitutes the core of learning analytics, is processed by machine learning algorithms to infer the learner's current state of knowledge and skill.[26] These inferences are used to populate and dynamically update a multidimensional student model, often represented as a competency map or knowledge graph, which provides a probabilistic estimate of the learner's proficiency on each competency in a framework.[28]

Table 1: Core Attributes of Learner-Centric Paradigms

Paradigm	Core Definition	Primary Goal	Unit of Progression	Primary Evidence
Competency-Based Education (CBE)	A systemic framework where progression is based on demonstrated mastery, making time variable.	Cultivation of transferable, applicable capabilities for real-world contexts.	Demonstrated competency, independent of seat time.	A body of evidence, often including performance on authentic tasks.
Mastery Learning	An instructional strategy requiring high proficiency on prerequisite units before advancing.	Ensuring foundational knowledge is secure and preventing cumulative learning gaps.	Mastery of a discrete instructional unit (e.g., >90% on a formative assessment).	Performance on criterion-referenced formative assessments.
Standards-Based Grading (SBG)	An assessment and reporting system that communicates achievement against specific learning standards.	Accurate, transparent, and equitable communication of what a student knows and can do.	N/A (It is a reporting method, not a progression model).	The most consistent and recent evidence of proficiency against a standard.
Personalized Learning	An instructional approach that optimizes pace, method, and content for each learner.	Maximized learner engagement, motivation, and efficiency by tailoring to individual needs.	Learner-driven pace through a customized pathway.	Continuous data on learner progress, engagement, and interaction patterns.

Evidence for this mechanism comes from its role as the foundational component of Intelligent Tutoring Systems (ITS), which have been the subject of decades of research. Multiple meta-analyses have established the effectiveness of ITS, which consistently produce significant learning gains over traditional classroom instruction.[29] Seminal reviews by Kulik & Fletcher, Koedinger et al., and VanLehn have shown that the real-time cognitive diagnosis provided by the student model is the critical element driving these positive outcomes.[31]

3.2 Adaptive Supply Pathway: The "Respond" Function

Once a diagnosis of the learner's state is established, the system must provide a targeted pedagogical response. This pathway automates the differentiation of instruction, a task that is exceptionally time-consuming for human teachers.

The primary mechanism here is the AI system reading from the student model to execute personalized instructional strategies. This includes adaptive sequencing, where the system alters the learner's path through the curriculum, presenting remedial content if a knowledge gap is detected or accelerating them to more advanced topics upon demonstrated mastery.[14] It also includes adaptive resource allocation, where algorithms recommend or deliver specific learning objects (e.g., videos, articles, practice problems) from a large content library that are best suited to address a specific learning need.[1] Finally, it can involve adaptive assessment, where the difficulty and focus of assessment items are dynamically adjusted based on the learner's real-time performance to more efficiently and accurately gauge their proficiency level.[14]

The effectiveness of these adaptive mechanisms is a core finding in the ITS literature. VanLehn's comprehensive 2011 review, for instance, found that ITS, which rely heavily on adaptive sequencing and feedback, achieve an effect size (d=0.76) nearly equivalent to that of one-on-one human tutoring (d=0.79), demonstrating the power of these automated responses.[33] A more recent meta-analysis confirmed that AI-assisted personalized learning has moderately positive effects on student outcomes, with intelligent feedback and learning path recommendation being particularly effective components.[35]

3.3 Teacher Orchestration Pathway: The "Amplify" Function

This pathway clarifies a critical aspect of the model: the AI does not replace the human teacher but rather augments their professional practice, amplifying their expertise and enabling them to focus on high-impact, uniquely human interactions.

The mechanism for this pathway involves the AI system aggregating and synthesizing data from the individual student models of an entire class. It then presents this information to the teacher through intuitive, actionable learning analytics dashboards.[36] These dashboards do not simply display raw data; they provide pedagogical insights. They can automatically flag students who are falling behind, identify common misconceptions across a cohort, suggest optimal groupings for collaborative projects based on complementary skill profiles, and provide evidence-based recommendations for targeted, small-group interventions.[1] The teacher uses these insights to orchestrate the complex social and pedagogical dynamics of the classroom, making data-informed decisions about where to focus their time and attention.

Evidence for this human-AI collaborative model shows a redefinition of the teacher's role from a "knowledge transmitter" to a "learning architect" or "meta-cognitive orchestrator".[1] While the technology provides the diagnostic and analytical heavy lifting, the teacher provides the crucial pedagogical judgment, socio-emotional support, and motivational guidance that are essential for deep learning.[37] The coherence and efficacy of these three pathways depend entirely on their ability to communicate through a common, robust data structure. The Shared Student Model is this central coupling mechanism. The Diagnostic Pathway continuously writes to this model. The Adaptive Supply Pathway reads from this model to deliver personalized content to the student. The Teacher Orchestration Pathway reads from and visualizes this model to provide actionable insights to the teacher. This architecture transforms three separate AI functions into a single, integrated cognitive system and underscores the critical importance of technical interoperability, which will be discussed in Section 6.

3.4 Testable Propositions

This mechanism model is not merely a conceptual framework; it generates a series of falsifiable hypotheses that can and should be tested empirically. The following propositions provide a concrete agenda for future research: (H1) Adaptive sequencing, informed by a dynamic student model, reduces the time required to remediate knowledge gaps, thereby increasing the rate of competency mastery compared to a fixed curriculum. (H2) The efficacy of teacher interventions prompted by an AI dashboard is moderated by the teacher's data literacy; above a certain literacy threshold, intervention frequency will positively correlate with learning gains for at-risk students. (H3) The consistency and portability of a shared student model across multiple learning tools will positively correlate with the student's ability to transfer learned competencies to novel, cross-domain tasks. (H4) Students whose learning pathways are governed by an "Adaptive-Autonomy Curve" will demonstrate significantly greater gains on standard measures of self-regulated learning (SRL) compared to students in a purely adaptive system. (H5) Performance-based assessments that incorporate process-level trace data (e.g., strategy shifts, error correction latency) will be more predictive of future transfer performance than assessments based solely on final outcomes. (H6) The implementation of systematic bias audits and human-in-the-loop (HITL) protocols for high-stakes AI-driven recommendations will significantly reduce performance gaps between demographic subgroups.

4 Re-architecting Pedagogy for Human-AI Collaboration

The integration of AI into a CBE framework necessitates a fundamental re-architecting of pedagogy. This section formalizes the collaborative relationship between the student, the human teacher, and the AI system using the theoretical lens of distributed cognition. It then introduces a novel design principle—the Adaptive-Autonomy Curve—to operationalize the balance between AI-driven personalization and the crucial development of learner autonomy.

4.1 A Distributed Cognition Framework

The theory of Distributed Cognition, pioneered by Edwin Hutchins, provides a powerful framework for understanding the AI-enabled classroom.[16] This theory posits that cognitive processes are not confined to the mind of a single individual but are distributed across a system of interacting agents and artifacts.[26] In this new educational ecosystem, the unit of analysis is not the individual learner, teacher, or AI tool, but the integrated cognitive system composed of all three. This framework allows for a strategic and complementary division of cognitive labor.

The role of the AI system is to manage a scale and complexity of information that is beyond human cognitive capacity. This includes the real-time tracking of learning trajectories for every student, the analysis of massive datasets to identify subtle patterns and misconceptions, the provision of instant, scalable feedback on well-defined tasks, and the logistical management of thousands of potential learning pathways.[32]

The role of the human teacher, liberated from the cognitive burdens of routine data collection, direct knowledge transmission, and administrative tasks, is elevated to that of a high-level learning architect and orchestrator.[1] The teacher focuses on the quintessentially human aspects of education: fostering creativity, critical thinking, and collaboration; providing nuanced, empathetic guidance and socio-emotional support; making complex pedagogical judgments that synthesize AI-provided data with their own professional wisdom; and designing the overarching learning experiences and culture of the classroom.[38]

The role of the student is no longer that of a passive recipient of information but an active agent and co-creator in their learning journey. They interact directly with the AI system (the "intelligent companion") for personalized practice and inquiry, collaborate with peers in AI-facilitated environments, and engage with the human teacher for higher-order mentorship and guidance.[38]

This division of labor is visualized in Figure 1, which models the primary information flows within this distributed cognitive triad. The model highlights the symbiotic relationship where each component performs the tasks for which it is best suited, creating a cognitive system more powerful than the sum of its parts.

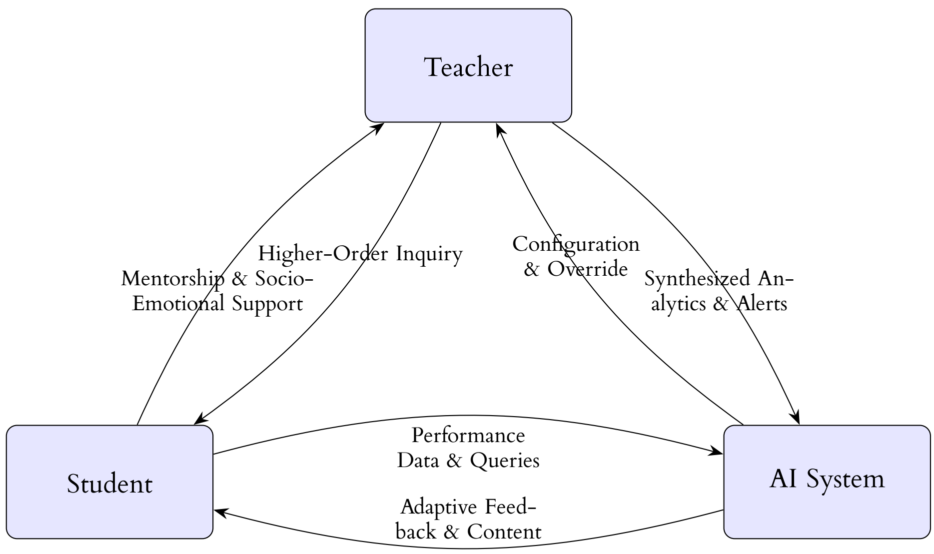

Figure 1: The Distributed Cognitive Triad in CBE

4.2 The Adaptive-Autonomy Curve: Balancing Personalization and Self-Regulation

A critical tension exists within this new pedagogical model. An overly prescriptive adaptive system, while optimizing for immediate content mastery, risks undermining the development of Self-Regulated Learning (SRL)—the proactive process through which students manage their own learning by setting goals, selecting strategies, and monitoring their progress.[24] SRL is a cornerstone of lifelong learning, and fostering it is a primary goal of education.[40] If the AI makes all the decisions, the learner may develop a form of cognitive dependency, becoming a passive follower of an optimized path rather than an autonomous agent of their own learning.[1]

To resolve this tension, this paper proposes the Adaptive-Autonomy Curve, a design principle for AIEd systems based on the concept of scaffolded release of control. The goal is to adapt not only the difficulty of the content but also the level of autonomy afforded to the learner, gradually ceding control as the student demonstrates greater self-regulatory capacity. The system begins with a high degree of structure and scaffolding, with the AI making most decisions about the learning path. As the student demonstrates both content mastery and behaviors indicative of SRL, the system progressively increases the learner's agency. This progression is not linear but follows a curve, potentially with periods of increased scaffolding if a student begins to struggle in a more autonomous environment.

The system's decision to release control is not arbitrary but is based on observable proxies for SRL derived from trace data. Drawing on the extensive literature on measuring SRL, these can include metrics related to *planning/forethought*, such as the frequency of accessing learning objective pages, use of in-system goal-setting tools, and time spent reviewing instructions before starting a task; metrics for *performance/monitoring*, such as the effective use of help-seeking features (e.g., using hints progressively vs. immediately revealing the answer), patterns of self-correction after errors, and time management across tasks; and indicators of *reflection*, such as the use of self-assessment rubrics, frequency of reviewing past performance on dashboards, and annotating or flagging content for later review.[40]

This principle is visualized in Figure 2. The curve represents a pedagogical strategy for teachers and system designers to intentionally manage the transition from a highly supported learning environment to one that fosters independent, lifelong learners.

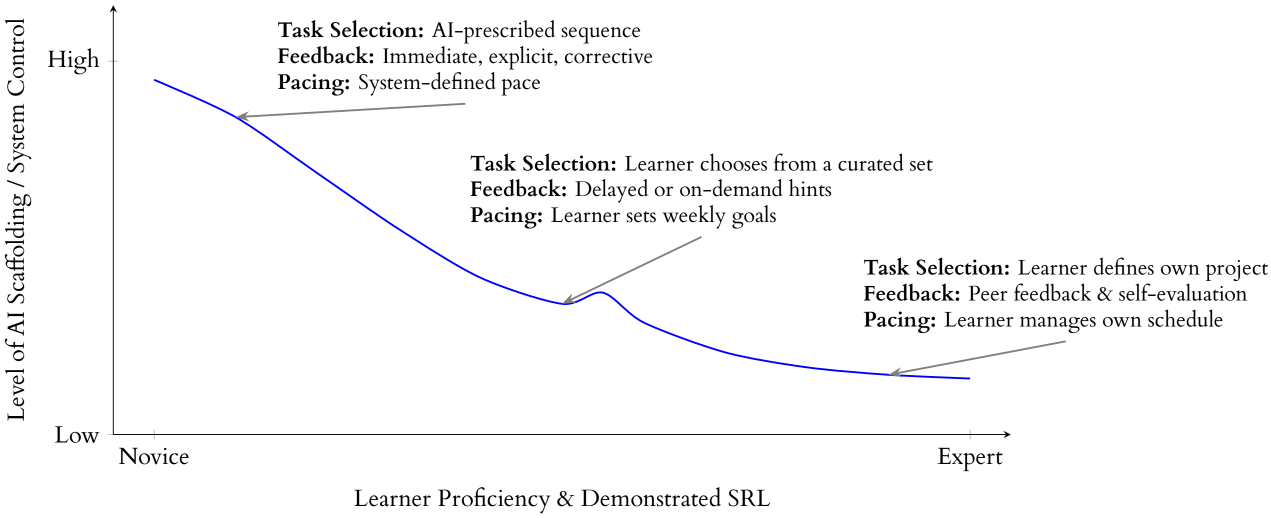

Figure 2: The Adaptive-Autonomy Curve

5 From Tasks to Transfer: Situated and Performance-Based Assessment

The cultivation of deep, transferable competencies, as defined by the OECD, demands an assessment paradigm that moves beyond measuring knowledge recall to evaluating the application of skills in complex contexts.[23] Traditional scalable assessments, such as multiple-choice tests, are fundamentally misaligned with the goals of CBE. The same AI technologies that enable personalized pedagogy also provide the means to create authentic, scalable, performance-based assessments that can resolve this historical disconnect. This section proposes a formal, multi-stage pipeline for this new mode of evaluation.

5.1 The Situated Performance-Based Assessment Pipeline

This pipeline outlines a systematic process for transforming the rich data generated during complex, interactive learning experiences into valid and reliable evidence of competency. It is grounded in the principles of "stealth assessment," where assessment is seamlessly embedded within the learning activity itself, reducing test anxiety and capturing a more authentic picture of student capabilities.[27]

The first stage is *Event Log & Trace Data Capture*. The process begins within an intelligent, situational learning environment, such as a scientific simulation, a historical role-playing game, or a complex design challenge.[1] As the learner engages with the task, the system captures a high-fidelity stream of interaction data in log files. This includes every action, decision, tool used, and communication sent, creating a detailed record of the learner's problem-solving process.[27]

The second stage, *Process Indicator Extraction*, addresses the fact that raw log data is information-rich but semantically poor. In this stage, AI algorithms analyze the event stream to extract meaningful, evidence-based process indicators. These are observable behaviors that serve as proxies for the unobservable cognitive competencies being assessed.[45] Examples include identifying a problem-solving strategy, measuring persistence, detecting creativity, and calculating efficiency.

In the third stage, *Structured Rubric Application*, the extracted process indicators, along with the final outcome of the task, are used to automatically score a predefined, structured rubric. This rubric translates qualitative behaviors into quantitative or categorical judgments of proficiency. An example rubric fragment for the competency "Adaptive Problem-Solving" is provided in the table below.

Table 2: Example Rubric Fragment for "Adaptive Problem-Solving"

Level	Descriptor	Observable Indicators (from Trace Data)
1 (Novice)	Persists with a single, ineffective strategy despite repeated failure.	Repeated sequences of identical failed actions; high frequency of "bottom-out" hint usage.
2 (Developing)	Acknowledges failure and attempts a different strategy, but the new strategy is not well-chosen.	A clear shift in action patterns after >3 failures; new strategy is random or inefficient.
3 (Proficient)	Systematically analyzes failure and selects an appropriate alternative strategy.	Low error-correction latency; new strategy directly addresses the cause of the previous failure.
4 (Exemplary)	Not only adapts strategy but also re-frames the problem, leading to a more elegant or efficient solution.	Use of advanced or non-obvious tool combinations; solution path is significantly shorter than the standard path.

The fourth stage is *Human-AI Co-Assessment and Calibration*. While AI can handle the scoring of well-defined indicators at scale, human judgment remains essential for evaluating nuance, creativity, and complex reasoning, as well as for ensuring fairness.[46] A human-AI co-assessment model optimizes the use of teacher time and expertise. The AI provides an initial score and a confidence rating for all student submissions, then flags a subset for human review based on several priority rules: a high-stakes priority for all summative assessments, an uncertainty priority for submissions where the AI's confidence is low, an anomaly priority for statistical outliers, and a calibration priority using a random sample to monitor the AI's performance.

The fifth and final stage is *Bias Audit and Explainability Report*, which focuses on accountability and transparency. The system generates regular reports that include inter-rater reliability metrics between the AI and human graders. Crucially, it performs systematic bias audits by disaggregating the AI's scoring accuracy and error rates across different demographic subgroups of students. This allows the institution to proactively identify and mitigate algorithmic bias, ensuring that the assessment system is fair and equitable for all learners.[1] The system should also provide an explainability report for any given score, linking the rubric level back to the specific process indicators and event log data that generated it.

6 Systemic Governance and Technical Interoperability

The successful implementation of the proposed model is contingent not only on pedagogical and technological innovation but also on the establishment of robust frameworks for governance and technical interoperability. A piecemeal adoption of AI tools without systemic oversight risks amplifying inequity, compromising student privacy, and creating a fragmented, ineffective digital ecosystem.[1] This section presents an integrated framework for managing these challenges, mapping implementation strategies to established international standards. To synthesize the relationship between the challenges of CBE, the opportunities presented by AI, and the necessity of governance, Figure 2 presents a three-dimensional matrix. This model provides a comprehensive overview for institutional leaders, illustrating how specific AI affordances can address known CBE implementation barriers, but only when coupled with appropriate governance controls.

(Description of Figure 2) Figure 2: A Three-Dimensional Matrix of CBE Barriers, AI Affordances, and Governance Controls

This figure would be a multi-layered table or a conceptual 3D cube. Rows (CBE Implementation Barriers): This axis lists the primary challenges to scaling CBE, drawn from implementation research.[7] Examples include: (1) High teacher workload for personalization, (2) Difficulty in continuous progress tracking, (3) Lack of differentiated learning resources, (4) Challenges in scalable, authentic assessment. Columns (AI Affordances): This axis lists the corresponding AI capabilities that address each barrier. For example: (1) AI-driven adaptive learning pathways, (2) Automated student modeling and dashboards, (3) AI-powered content curation and generation, (4) The Situated Performance-Based Assessment Pipeline. Depth (Governance

& Control Layers): This axis lists the essential governance mechanisms required to manage the risks introduced by each AI affordance. For example: (1) For adaptive pathways: Algorithmic transparency and bias audits. (2) For student modeling: Data privacy policies (FERPA/GDPR) and security protocols. (3) For content generation: Policies on academic integrity and critical AI literacy training. (4) For AI-based assessment: Human-in-the-loop (HITL) review protocols and validation studies.

6.1 A Modular Governance Framework (Compliance Mapping)

To move from concept to practice, institutions need an actionable governance framework. The proposed modular structure maps key areas of responsibility to established international standards, providing a clear path toward responsible implementation.

The *Policy Layer (Data & Rights)* concerns the protection of student data and rights. All practices must be mapped to the requirements of relevant legal frameworks such as the Family Educational Rights and Privacy Act (FERPA) in the US and the General Data Protection Regulation (GDPR) in Europe. Core actions include implementing principles of data minimization (collecting only necessary data), purpose limitation (using data only for specified educational purposes), and establishing clear consent protocols for students and parents.[1]

The *Professional Development (PD) Layer (People & Pedagogy)* focuses on equipping educators with the necessary skills to operate effectively and ethically in an AI-rich environment. Professional development programs should be aligned with the UNESCO AI Competency Framework for Teachers, which outlines the knowledge, skills, and values teachers need.[36] Core actions include training on data literacy (how to interpret dashboards and make sound pedagogical judgments), ethical AI use (understanding potential biases and limitations), and new pedagogical strategies like orchestrating a distributed cognitive classroom.

The *Infrastructure Layer (Technology & Risk)* addresses the management of the AI systems themselves. The design, deployment, and evaluation of AI tools should be governed by the NIST AI Risk Management Framework (AI RMF), which provides a structured process to govern, map, measure, and manage AI risks.[47] For organizations seeking formal certification, establishing an AI Management System (AIMS) compliant with ISO/IEC 42001 is the goal. Core actions under this layer include conducting regular bias audits, establishing clear HITL workflows for high-stakes decisions, and implementing rigorous vendor assessment protocols to ensure third-party tools meet institutional standards.[49]

The *Evaluation Layer (Efficacy & Equity)* ensures a continuous cycle of review and improvement. The institution must establish processes for the independent and ongoing evaluation of its AI systems. Core actions include conducting A/B tests to assess the pedagogical efficacy of algorithmic changes and generating regular reports on system performance disaggregated by student subgroups to monitor for and address any emergent equity gaps.

6.2 Beyond LTI: The Need for Pedagogical Interoperability

The distributed, multi-tool ecosystem envisioned in this paper requires a new level of technical interoperability. Current standards like Learning Tools Interoperability (LTI) and OneRoster are essential for administrative interoperability—they allow different systems to exchange basic data like class rosters and final grades.[51] However, they are insufficient for the deep pedagogical integration required by our model. They lack the vocabulary and structure to share the rich, dynamic state of a student's competency map, their learning process data, or their affective state.[52] Without a shared pedagogical language, the digital learning environment remains a collection of disconnected silos.

To address this, we propose the development of a Shared Student Model as a new interoperability protocol. This protocol would define a standard, machine-readable format for representing and exchanging a student's evolving competency profile. This model would serve as the "lingua franca" that allows a diverse suite of AI tools—an ITS for math, a writing feedback agent, a science simulation—to both contribute to and draw from a single, coherent, and portable understanding of the learner. To make this proposal concrete, Table 2 outlines a minimal viable set of data fields for such a model. This draft specification draws on the strengths of existing 1EdTech (now 1EdTech) standards: it uses the Competencies and Academic Standards Exchange (CASE) standard for uniquely identifying competencies and the Caliper Analytics standard for linking to the underlying evidence streams.[54]

7 Discussion and Boundary Conditions

The integrated framework proposed in this paper—encompassing a testable mechanism model, a re-architected pedagogy, a performance-based assessment pipeline, and a multi-layered governance structure—offers a comprehensive

Table 3: Data Field Specifications for a Shared Student Model

Field Name	**Data Type**	**Description**	**Rationale / Standard Link**
`competency_id`	URN/GUID	A globally unique identifier for the competency node.	Aligns with 1EdTech CASE for referencing a specific standard in a published framework.
`mastery_level`	Integer/Enum	The student's current estimated proficiency level on a defined scale (e.g., 1-4).	The core state variable for CBE progression.
`confidence_score`	Float (0.0-1.0)	The system's probabilistic confidence in the `mastery_level` estimate.	Essential for human-AI co-assessment and identifying areas of uncertainty.
`latest_evidence_event`	URN/GUID	The unique identifier of the most recent learning event that updated the model's state.	Provides traceability and links to the 1EdTech Caliper event that served as evidence.
`evidence_stream_url`	URL	A link to the endpoint for the full stream of Caliper events associated with this competency.	Enables deep-dive analysis and auditing of the evidence base for a given competency claim.
`affective_motive_signal`	Enum	An optional field for signals related to engagement, frustration, or motivation (e.g., 'Engaged', 'Struggling').	Allows systems to adapt not just to cognitive state but also to affective state, supporting SRL.
`last_updated_timestamp`	ISO 8601 DateTime	The timestamp of the last update to this competency node.	Critical for synchronization and understanding the recency of the evidence.

roadmap for realizing the potential of AI-scaled Competency-Based Education. However, the successful application of this framework is not universal; it is subject to a set of critical boundary conditions and limitations that must be carefully considered by researchers, policymakers, and practitioners.

7.1 Boundary Conditions of the Model

The efficacy and feasibility of this model are contingent upon several contextual factors.

The model's efficacy is contingent on *disciplinary differences*. It is most readily applicable in well-structured domains such as mathematics, computer science, and certain areas of the natural sciences. In these fields, competencies can often be decomposed into hierarchical knowledge components and procedural skills that are amenable to automated assessment and adaptive sequencing. The model faces greater challenges in ill-structured domains like the humanities, social sciences, and the arts. In these areas, competencies such as "historical interpretation," "ethical reasoning," or "aesthetic judgment" are less decomposable and their assessment is inherently more interpretive and context-dependent. Implementing the model in such domains would require a significantly greater reliance on the human-AI co-assessment stage, with the AI's role shifting from automated scoring to primarily organizing evidence and highlighting key passages for human evaluation.

A second condition is the *teacher data literacy threshold*. The Teacher Orchestration Pathway is the lynchpin that connects the AI's analytical power to meaningful classroom practice. However, its effectiveness is entirely dependent on the teacher's ability to interpret the data presented on analytics dashboards and translate those insights into sound pedagogical decisions. This implies a necessary threshold of teacher data literacy. Below this threshold, dashboards risk becoming a source of confusion or may be ignored, rendering the orchestration pathway inert. This has profound implications for professional development, which must evolve beyond simple "tool training" to cultivate a deeper capacity for data-informed pedagogical reasoning, ethical reflection on algorithmic recommendations, and the systemic thinking required to manage a distributed cognitive classroom.

Finally, the model is constrained by *infrastructural and economic barriers*. The implementation of this model presupposes a robust technological infrastructure, including reliable high-speed internet access, adequate computing devices for all students, and the institutional capacity to procure, integrate, and maintain sophisticated AI systems.[1] This raises significant equity concerns. Without deliberate and substantial public investment to ensure this infrastructure is available to all schools, not just the most affluent, the adoption of such advanced models could dramatically widen the existing digital divide, creating a new chasm between students who benefit from AI-scaled CBE and those who remain in traditional systems.[23] The economic costs of both the technology and the requisite comprehensive professional development represent a major barrier that must be addressed at the policy level.

7.2 Limitations and Future Directions

This paper presents a theoretical model and an implementation framework. While grounded in extensive empirical literature on its component parts (e.g., ITS, SRL, learning analytics), the integrated model itself requires comprehensive empirical validation. The testable propositions offered in Section 3 provide a starting point for this work, but the complex, systemic nature of the framework calls for a multi-faceted research agenda. Longitudinal, design-based research studies are needed to understand how these components interact in authentic classroom settings over time. Furthermore, this paper has focused primarily on the cognitive and pedagogical dimensions of learning. Future work should more deeply integrate the socio-emotional aspects, exploring how the human-AI collaborative model can be optimized to foster not only competency but also student well-being, sense of belonging, and interpersonal skills. Finally, the proposal for a "Shared Student Model" is conceptual; its translation into a robust, secure, and widely adopted technical standard will require a significant collaborative effort among researchers, standards bodies like 1EdTech, and technology developers.

8 Conclusion and Future Research

This paper has articulated a comprehensive framework for the systemic integration of Artificial Intelligence into a Competency-Based Education paradigm. By moving beyond a descriptive account of AI as a tool, it has proposed a testable mechanism model that specifies how AI can resolve the historical scalability challenges of CBE. The framework's core contributions lie in its conceptual rigor, its formalization of a human-AI collaborative pedagogy through the lens of distributed cognition, and its operationalization of key design principles like the Adaptive-Autonomy Curve and the Situated Performance-Based Assessment Pipeline. Crucially, it embeds these innovations within a robust structure for governance and technical interoperability, providing a pragmatic roadmap for responsible implementation.

The synthesis of pedagogical theory, cognitive science, and technical standards presented here offers a new lens through which to view the future of AI in education. The ultimate goal is not the automation of teaching but the augmentation of the entire educational ecosystem, creating a system where technology manages complexity at scale, freeing human educators to focus on the uniquely human dimensions of learning. This vision shifts the focus from teaching with AI to designing learning environments that cultivate the competencies needed to thrive in the age of AI.

To transform this theoretical framework into evidence-based practice, a concerted and rigorous research agenda is required. The propositions (H1-H6) advanced in this paper provide a direct path for this empirical work. The following outlines a concrete agenda for future research.

Regarding *experimental and quasi-experimental designs*, the causal claims embedded in the propositions should be tested through controlled studies. For instance, to test H1 (Diagnosis $\rightarrow$ Mastery Rate) and H2 (Orchestration $\times$ Teacher Literacy $\rightarrow$ Intervention Effect), a cluster-randomized trial could be conducted. Schools could be randomly assigned to one of three conditions: (1) a traditional curriculum, (2) a CBE curriculum with an AI platform providing only the diagnostic and adaptive supply pathways for students, and (3) the full model, including the teacher orchestration dashboard. Student outcomes would be measured by mastery rate and learning gains, while teacher data literacy would be measured as a potential moderating variable. To test H4 (SRL $\rightarrow$ Adaptive-Autonomy Curve), a within-subjects or between-subjects experiment could compare two versions of an adaptive learning system: one that is purely adaptive versus one that implements the Adaptive-Autonomy Curve. Dependent variables would include not only performance on the learning tasks but also validated SRL measures (e.g., questionnaires like the MSLQ) and behavioral trace data proxies for SRL.

In the area of *design-based research and validation studies*, the novel constructs proposed require iterative development and validation in authentic contexts. The Situated Performance-Based Assessment Pipeline (H5) should be developed through a design-based research methodology. Researchers would partner with educators to co-design and refine the pipeline within a specific course, iteratively improving the process indicators, rubrics, and human-AI co-assessment workflow while collecting validity evidence by correlating the pipeline's outputs with external measures of student performance and transfer. Similarly, the Shared Student Model (H3) requires a multi-institutional pilot study. A consortium of institutions and technology vendors could collaborate to implement a prototype of the standard, tracking student data as they move between different interoperable tools. The primary outcome would be the correlation between the consistency of the student model across platforms and performance on capstone projects that require the transfer of skills learned in different tools.

Finally, the *development of a common metric and indicator set* is crucial as the field currently lacks standardized, validated metrics for many of the key processes described in this paper. A critical line of future work is the development of a shared repository of operationalized metrics derived from trace data. This would include defining and validating measures for constructs such as "mastery velocity" (the rate of change in a student's competency model), "error remediation efficiency" (the time and number of attempts required to correct a misconception after feedback), and a suite of "SRL-in-action" indicators. The creation of such a shared set of metrics would be invaluable for enabling cumulative, comparable, and replicable research across the field, accelerating our collective understanding of how to design truly effective and equitable human-AI orchestrated classrooms.

To Cite This Article Xin HUANG. (2025). Designing Human-AI Orchestrated Classrooms: Mechanisms, Protocols, and Governance for Competency-Based Education. *Artificial Intelligence Education Studies*, *1*(3), 14–27. https://doi.org/10.6914/aiese.010302

Reference

[1] Dede, C. (2017). A 21st century education framework for AI-enhanced personalized learning. In M. A. Peters (Ed.), *Encyclopedia of educational philosophy and theory* (pp. 1 6). Springer.

[2] Tennessee Comptroller of the Treasury, Office of Research and Education Accountability. (2016, July). *Competency-based Education in Grades K-12*. https://comptroller.tn.gov/content/dam/cot/orea/advanced-search/2016/2016_OREA_CompetencyBasedEd.pdf

[3] Aurora Institute. (n.d.). *Competency-based education*. Retrieved September 2, 2025, from https://aurora-institute.org/our-work/competencyworks/competency-based-education/

[4] American Association of Colleges of Nursing. (n.d.). *What is competency-based education?*. Retrieved September 2, 2025, from https://www.aacnnursing.org/essentials/tool-kit/competency-based-education

[5] Austin Community College District. (n.d.). *What is Competency-based Education (CBE)?*. Retrieved September 2, 2025, from https://sites.austincc.edu/cbec/what-is-cbe/

[6] United Nations Educational, Scientific and Cultural Organization. (2018). *UNESCO ICT competency framework for teachers*. https://unesdoc.unesco.org/ark:/48223/pf0000265721

[7] The Core Collaborative. (2024). *Navigating the pitfalls of competency-based learning*. Retrieved September 2, 2025, from https://thecorecollaborative.com/navigating-the-pitfalls-of-competency-based-learning/

[8] Newlane University. (n.d.). *Tackling challenges while adopting competency-based education*. Retrieved September 2, 2025, from https://newlane.edu/challenges-while-adopting-competency-based-education-2/

[9] Sturgis, C. (2016, October 19). *Mapping out the terrain of competency-based education implementation*. Aurora Institute. https://aurora-institute.org/cw_post/mapping-out-the-terrain-of-competency-based-education-implementation/

[10] U.S. Department of Education. (n.d.). *Direct assessment (competency-based) programs*. Retrieved September 2, 2025, from https://www.ed.gov/laws-and-policy/higher-education-laws-and-policy/higher-education-policy/direct-assessment-competency-based-programs

[11] Hall, B. (2024, January 22). *Mastery learning: A mastery approach to learning*. Structural Learning. https://www.structural-learning.com/post/mastery-learning

[12] Wikipedia contributors. (2025, August 24). Mastery learning. In *Wikipedia, The Free Encyclopedia*. Retrieved September 2, 2025, from https://en.wikipedia.org/wiki/Mastery_learning

[13] Carlton, G. (2024, April 11). *What is mastery learning model? Definition, principles, and examples for 2025*. Research.com. https://research.com/education/what-is-mastery-learning

[14] University of California, Irvine, Division of Teaching Excellence and Innovation. (n.d.). *Mastery learning*. Teaching Experiment Academy. Retrieved September 2, 2025, from https://tea.dtei.uci.edu/resources/mastery-learning/

[15] Lee, W.-Y., & Tso, W. W. S. (2023). A practical review of mastery learning. *Medical Education Online*, *28*(1), 2209017. https://doi.org/10.1080/10872981.2023.2209017

[16] Choppin, J. D. (2023, November). *Standards-Based Grading: History, Practices, Benefits, Challenges, and Next Steps*. Center for Urban Education Success, University of Rochester. https://www.rochester.edu/warner/cues/wp-content/uploads/2023/11/Standards-Based-Grading.pdf

[17] Guskey, T. R., & McTighe, J. (2024, May). Standards-based grading: Core principles and current implementation debates. *Principal Leadership*, *24*(9), 42 47.

[18, 19] Wikipedia contributors. (2025, August 23). Personalized learning. In *Wikipedia, The Free Encyclopedia*. Retrieved September 2, 2025, from https://en.wikipedia.org/wiki/Personalized_learning

[20] Ohio Department of Education and Workforce. (n.d.). *Personalized learning*. Retrieved September 2, 2025, from https://education.ohio.gov/Topics/Teaching/Personalized-Learning

[21] Ohio Department of Education. (n.d.). *Personalized learning frequently asked questions*. Retrieved September 2, 2025, from https://education.ohio.gov/Topics/Teaching/Personalized-Learning/Resources/Personalized-Learning-Frequently-Asked-Questions

[22] American Public University. (2024, January 26). *What is personalized learning? How does it impact students?*. https://www.apu.apus.edu/area-of-study/education/resources/what-is-personalized-learning/

[23] Organisation for Economic Co-operation and Development. (2010). *The definition and selection of key competencies: Executive summary*. https://one.oecd.org/document/EDU/EDPC/ECEC/RD(2010)26/en/pdf

[24] United Nations Educational, Scientific and Cultural Organization. (2017). *UNESCO Competency Framework*. https://en.unesco.org/sites/default/files/competency_framework_e.pdf

[25] UNESCO International Institute for Higher Education in Latin America and the Caribbean. (2016). *Competency framework for university professionals*. https://unesdoc.unesco.org/ark:/48223/pf0000245056

[26] Dillenbourg, P. (1996). *Some technical implications of distributed cognition on the design of interactive learning environments* [Preprint]. ResearchGate. https://www.researchgate.net/publication/243772698_Some_technical_implications_of_distributed_cognition_on_the_design_of_interactive_learning_environments

[27] Shute, V. (2011). *Stealth assessment* (Report No. ED612156). ERIC. https://files.eric.ed.gov/fulltext/ED612156.pdf

[28] Siemens, G., & Gašević, D. (2012). Learning analytics and educational data mining: Towards communication and collaboration. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 252 254). Association for Computing Machinery. https://doi.org/10.1145/2330601.2330661

[29] Ma, W., Adesope, O. O., Nesbit, J. C., & Liu, Q. (2014). Effectiveness of intelligent tutoring systems: A meta-analysis. *Review of Educational Research*, *84*(4), 548 583. https://doi.org/10.3102/0034654314548309

[30] International Society of the Learning Sciences. (n.d.). *Cognitive tutors*. Retrieved September 2, 2025, from https://www.isls.org/research-topics/cognitive-tutors/

[31] Paletz, S. B. F., D' Amico, J. A., Boylan, E. K., & Maxon, J. G. F. (2017, October). *Effectiveness of intelligent tutoring systems: A meta-analytic review* (Report No. P-8686). Institute for Defense Analyses. https://www.ida.org/-/media/feature/publications/w/we/welch-award-2017---effectiveness-of-intelligent-tutoring-systems-a-meta-analytic-review/1-effectivenessits.ashx

[32] Roll, I., & Chang, Y.-J. (2023, March 15). *Human-AI partnerships in education: Entering the age of collaborative intelligence*. The Learning Agency Lab. https://the-learning-agency.com/the-cutting-ed/article/human-ai-partnerships-in-education-entering-the-age-of-collaborative-intelligence/

[33, 34] VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, *46*(4), 197 221. https://doi.org/10.1080/00461520.2011.611369

[35] Tushabe, A. (2024, April 11). *A complete guide to personalized learning in K-12 education (2025)*. PowerSchool. https://www.powerschool.com/blog/complete-guide-personalized-learning-k12-education/

[36] Miao, F., Holmes, W., Huang, R., & Zhang, H. (2021). *AI and education: Guidance for policy-makers*. United Nations Educational, Scientific and Cultural Organization. https://unesdoc.unesco.org/ark:/48223/pf0000376709

[37] Das, S. (2025, February 24). *A framework for human-centric AI-first teaching*. AACSB International. https://www.aacsb.edu/insights/articles/2025/02/a-framework-for-human-centric-ai-first-teaching

[38] Sonnek, A. J. (2023). *Standards-based grading and the future of high school student assessment* [Doctoral dissertation, Bethel University]. Spark. https://spark.bethel.edu/cgi/viewcontent.cgi?article=2092&context=etd
[39] Learning and the Adolescent Mind. (n.d.). *Barry Zimmerman*. Retrieved September 2, 2025, from http://learningandtheadolescentmind.org/people_04.html
[40] Pu, T.-S., Chiu, M.-J., Cheng, S.-F., & Huang, I-W. (2017). A review of self-regulated learning: Six models and four directions for medical education. *Frontiers in Psychology*, *8*, 422. https://doi.org/10.3389/fpsyg.2017.00422
[41] Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice*, *41*(2), 64 70. https://doi.org/10.1207/s15430421tip4102_2
[42] Azevedo, R., Wiedbusch, M., & Eguchi, A. T. (2022). Self-regulated learning and learning analytics. In *Handbook of Learning Analytics* (2nd ed.). Society for Learning Analytics Research. https://doi.org/10.18608/hla22.008
[43] Perry, N. E., VandeKamp, K. V., Mercer, L. R., & Nordby, C. J. (2001, April). *Measuring self-regulated learning* [Paper presentation]. Annual Meeting of the American Educational Research Association, Seattle, WA, United States.
[44] Shute, V. J. (2022). *Stealth assessment: Measuring and supporting learning in video games*. The MIT Press. https://doi.org/10.7551/mitpress/14488.001.0001
[45] Google Cloud. (n.d.). *Run a computation-based evaluation pipeline*. Retrieved September 2, 2025, from https://cloud.google.com/vertex-ai/generative-ai/docs/models/computation-based-eval-pipeline
[46] Lee, D., Chilton, M. A., Smith, D. B. D., & Kim, H. W. H. (2021). Human-AI co-creation: A study of AI-assisted news writing. In *CHI Conference on Human Factors in Computing Systems Proceedings*. Association for Computing Machinery. https://doi.org/10.1145/3411764.3445657
[47] Skadden, Arps, Slate, Meagher & Flom LLP and Affiliates. (2023, May). *Evaluating and managing AI risk using the NIST framework*. https://www.skadden.com/insights/publications/2023/05/evaluating-and-managing-ai-risk-using-the-nist-framework
[48] Palo Alto Networks. (n.d.). *What is the NIST AI risk management framework?*. Retrieved September 2, 2025, from https://www.paloaltonetworks.com/cyberpedia/nist-ai-risk-management-framework
[49] Microsoft. (2024, February 1). *ISO/IEC 42001:2023 Artificial Intelligence management system*. Microsoft Learn. https://learn.microsoft.com/en-us/compliance/regulatory/offering-iso-42001
[50] ISMS.online. (n.d.). *Understanding ISO 42001 and AIMS*. Retrieved September 2, 2025, from https://www.isms.online/iso-42001/
[51] Consortium for School Networking. (n.d.). *K-12 technology interoperability standards & tools*. Retrieved September 2, 2025, from https://www.cosn.org/tools-resources/toolkits/interoperability/
[52] Wikipedia contributors. (2024, June 20). Schools Interoperability Framework. In *Wikipedia, The Free Encyclopedia*. Retrieved September 2, 2025, from https://en.wikipedia.org/wiki/Schools_Interoperability_Framework
[53] EdSurge. (2022). *What does it take to make interoperability work in K-12 education?*. https://www.edsurge.com/research/reports/what-does-it-take-to-make-interoperability-work-in-k-12-education
[54] 1EdTech. (n.d.). *Caliper metric profiles: Common explanations*. Retrieved September 2, 2025, from https://www.1edtech.org/specs/caliper/caliper-metric-profiles-common-explanations
[55] 1EdTech. (n.d.). *Caliper Analytics®*. Retrieved September 2, 2025, from https://www.1edtech.org/standards/caliper
[56] 1EdTech. (n.d.). *Competencies and Academic Standards Exchange®*. Retrieved September 2, 2025, from https://www.1edtech.org/standards/case
[57] 1EdTech. (n.d.). *Introduction to the 1EdTech CASE Standard*. Retrieved September 2, 2025, from https://www.1edtech.org/standards/case/about

Editor *Changkui LI wtocom@gmail.com*