Metaverse 2026, 7(2), 8542.
https://doi.org/10.54517/m8542

Review

The Educational Metaverse: A Narrative Review of Theoretical Frameworks, Learning Mechanisms, and Illustrative Cases from China

Dongxing Yu

Institute for Sustainable Development in Education, School of Education, Sanda University, Shanghai, China
*** Corresponding author:** Dongxing Yu, yudongxing@sandau.edu.cn

CITATION

Yu DX. The Educational Metaverse: A Narrative Review of Theoretical Frameworks, Learning Mechanisms, and Illustrative Cases from China. Metaverse. 2026; 7(2): 8542. https://doi.org/10.54517/m8542

ARTICLE INFO

Received: 19 March 2026
Revised: 8 April 2026
Accepted: 8 May 2026
Available online: 20 May 2026

Abstract: This narrative review synthesizes representative English- and Chinese-language scholarship, conference papers, policy documents, and institutional materials on the educational metaverse, drawn mainly from publications between 2021 and 2025 together with selected foundational sources. Rather than treating the field as a settled evidence base, the article offers a conceptual synthesis of three linked domains: theoretical foundations, learning and assessment mechanisms, and illustrative implementation cases from China. Grounded in constructivist pedagogy, embodied cognition theory, and dual-perspective design frameworks, the review examines the four-pillar typology of metaverse environments—Augmented Reality, Lifelogging, Mirror Worlds, and Virtual Reality—as well as the learning and evaluation logics associated with AI-enabled stealth assessment. China is used here as an important case for observing state-led educational metaverse deployment, including the MIIT 2023-2025 Action Plan, metaverse-engineering initiatives in higher education, vocational training experiments, and selected Metaverse Intelligent Education Technology (MIET) applications. Across these cases, the manuscript emphasizes reported tendencies, proposed architectures, and context-specific implementation lessons rather than universal causal claims. The review concludes by identifying major technical, socio-economic, and ethical challenges, especially the tension between rich biometric telemetry, equitable access, and privacy-centered governance.

Keywords: Educational Metaverse; Extended Reality (XR); Constructivism; Embodied Cognition; Stealth Assessment; Immersive Learning; Virtual Reality (VR); MIIT Action Plan; Higher Education; China; Metaverse Intelligent Education Technology (MIET).

1. Introduction

The educational metaverse is increasingly discussed as a shift from primarily screen-based learning toward more immersive, three-dimensional spatial-computing environments. Earlier phases of educational digitization—such as Web 2.0 platforms, video conferencing, and MOOCs—expanded access but were also associated in many settings with fatigue, uneven interaction, and low completion rates. The educational metaverse is best understood not as a settled replacement for these models, but as a family of XR-enabled environments that may expand interaction, co-presence, simulation, and practice under particular conditions [1,2].

Unlike conventional e-learning platforms that rely on abstraction and point-and-click interfaces, the educational metaverse is defined by its high degree of physical embodiment, spatial persistence, and synchronized, multi-user interaction. Within this emerging digital ecosystem, the learner's role shifts fundamentally from a passive observer to an active, physically engaged participant. Students enter these virtual environments via advanced smart wearables—such as head-mounted displays offering six-degrees-of-freedom tracking and spatial audio—and are represented by customizable avatars. This allows for real-time, multisensory engagement with three-dimensional scenes and dynamic learning resources. Instead of passively reading about a chemical reaction, learners can physically manipulate digital molecules; instead of watching a historical documentary, they can navigate a recreated ancient city alongside peers and intelligent pedagogical agents capable of providing real-time, conversational tutoring.

Fully leveraging this spatial technology may require substantial restructuring of pedagogical methodologies, assessment frameworks, and institutional infrastructures, shifting educational institutions from physical silos toward hybrid physical-digital ecosystems. This article therefore offers a narrative review and conceptual synthesis of the educational metaverse, tracing connections between learning theory, platform design, activity-embedded assessment logics such as stealth assessment, and implementation practice. It pays particular attention to China as a strategically important, but not statistically representative, case of state-led deployment. By examining policy documents, published studies, conference papers, and institutional examples together, the analysis clarifies how metaverse concepts are being translated into concrete educational experiments, Metaverse Intelligent Education Technology (MIET) applications, governance questions, workforce training agendas, and cultural-preservation projects.

1.1 Review scope and synthesis approach

This article is framed as a narrative review rather than a systematic review. To improve transparency, the review followed a purposive multi-source search and screening strategy conducted across Scopus, Web of Science, Google Scholar, CNKI, publisher databases, and official policy repositories. English- and Chinese-language searches used combinations of terms such as "educational metaverse", "edu-metaverse", "immersive learning", "virtual reality", "augmented reality", "stealth assessment", "AI tutoring", "metaverse engineering", "China", as well as Chinese equivalents including "教育元宇宙", "元宇宙教育", and "沉浸式学习". Sources were included when they addressed education or training rather than entertainment alone, contributed directly to one of four analytic themes (metaverse typologies, learning-theory foundations, assessment and AI adaptation, and China-based implementation), and offered conceptual, empirical, policy, or implementation relevance. Preference was given to peer-reviewed journal articles and scholarly book chapters when interpreting learning outcomes. Conference papers, policy documents, interviews, and institutional materials were retained mainly when they documented emerging architectures, initiatives, or governance frameworks not yet well covered in peer-reviewed literature. The core corpus is concentrated in the 2021–2025 period, supplemented by selected foundational works where conceptually necessary. China is used here as an illustrative leading case of state-coordinated deployment rather than as a statistically representative sample of global practice.

To make the scale of the review more concrete, the final narrative synthesis in this article is built around 20 directly cited core sources, selected from a broader purposive screening of English- and Chinese-language candidate materials across the databases and repositories noted above. Materials were excluded from the core synthesis when they focused primarily on entertainment rather than education or training, provided only promotional discussion without substantive analytical or implementation detail, duplicated claims already represented by stronger sources, or lacked sufficient relevance to one of the article's four analytic themes.

Although China is the main empirical focus of the implementation discussion, related educational-metaverse experimentation has also been reported in other regional contexts, including South Korea, Europe, and North America, typically through institution-level pilots and conceptual frameworks rather than a single coordinated national model [1,2].

2. Theoretical Architecture of the Educational Metaverse

The educational metaverse is better understood as a layered socio-technical assemblage than as a single technology. A useful review therefore needs to connect computational infrastructure, interaction design, and learning-science concepts while remaining explicit about differences in evidentiary strength across claims.

2.1 The Four-Pillar Typology of Metaverse Environments

To map technological capabilities to pedagogical needs, educational-metaverse scholarship often distinguishes among four broad modalities along the axes of augmentation versus simulation and internal identity versus external environment [1,2]. These categories should be treated as analytical lenses rather than as rigid or mutually exclusive system types:

- Augmented Reality (AR): Superimposes digital information onto the physical world. Typical educational affordance: can support situated learning by reducing split attention between a physical task and explanatory information.
- Lifelogging: The continuous capture of everyday experiences and physiological or behavioral data. Typical educational affordance: can support longitudinal observation of participation, reflection, or skill development, although it also raises substantial privacy questions.
- Mirror Worlds: Persistent digital replicas (digital twins) of external environments. Typical educational affordance: can support systemic reasoning, scenario testing, and simulation of complex real-world settings.
- Virtual Reality (VR): Fully synthetic digital environments. Typical educational affordance: can increase presence and enable practice with abstract, distant, or hazardous phenomena, while also carrying risks of overload or simulator sickness when poorly designed.

2.2 Constructivism and Embodied Cognition

The theoretical rationale for metaverse-based learning is often drawn from constructivism and embodied cognition. Constructivist perspectives emphasize active meaning-making and social scaffolding, whereas embodied-cognition accounts suggest that sensorimotor interaction can shape understanding and recall. These theories provide a useful rationale for spatial learning design, but they should not be treated as guarantees that immersive environments will outperform conventional instruction in every context.

In practice, immersive environments may support learning by aligning conceptual exploration with movement, gesture, and spatial orientation. For example, turning to inspect a 3D model or manipulating virtual objects can distribute cognition across visual, motor, and spatial channels. At the same time, these potential benefits depend on design quality, learner familiarity, and task fit; poorly designed immersion can also increase extraneous load.

The concept of the "Edu-Metaverse of Knowledge" offers one concrete example of this design logic [3]. In that framework, academic curricula are organized as navigable three-dimensional knowledge structures so that learners move between conceptual nodes and relationships through spatial exploration.

Illustrative empirical evidence comes from the "K-Cube VR" prototype reported by Sin et al. [3]. In a comparative evaluation of conceptual retention, students using

K-Cube VR obtained an intersection over union (IoU) accuracy score of 59.81%, compared with 55.19% for a 2D desktop-interface cohort and 39.70% for a traditional control group. These results are better interpreted as preliminary support for spatially organized interaction than as definitive validation of embodied cognition in educational metaverse design.

2.3 The Dual-Perspective Architecture: Learners and Educators

The long-term scalability and pedagogical effectiveness of the educational metaverse require a highly structured, dual-perspective framework. Historically, ed-tech interventions have faltered by over-indexing on student engagement while ignoring the immense logistical and operational burdens placed on teachers. A successful spatial computing ecosystem must simultaneously optimize the cognitive environment for the learner and the orchestrational environment for the educator.

• Learner Perspective and Cognitive Optimization: From a learner perspective, the Cognitive Affective Model of Immersive Learning (CAMIL) is useful because it treats immersion as conditionally beneficial rather than inherently superior. High fidelity may increase presence and engagement, but it can also overwhelm working memory or contribute to simulator sickness when locomotion and interface design are poorly aligned with the task. For that reason, interface choices such as simplified locomotion, diegetic tools embedded in the virtual world, and gradual task escalation should be understood as design strategies for managing cognitive load, not as universally optimal solutions.

• Educator Perspective and Generative Orchestration: From an educator perspective, immersive platforms also raise orchestration demands. Teachers must manage avatar activity, social interaction, task transitions, and technical contingencies, often at the same time. AI-supported pedagogical agents may reduce some of this burden by handling routine questions, role-play, or adaptive hinting in sparsely populated environments, but they do not remove the need for human instructional judgment [1]. The layered synthesis used in this review is summarized in **Figure 1**.

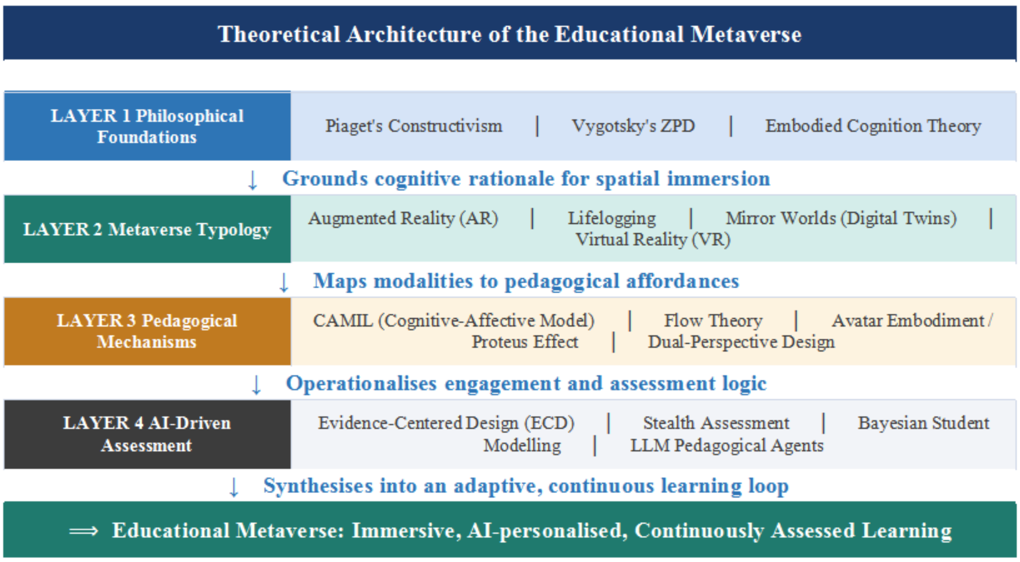

Figure 1. Theoretical Architecture of the Educational Metaverse: A four-layer progression from foundational learning science to AI-driven spatial learning systems

Note: Each layer is intended to inform the next. Philosophical foundations (Layer 1) guide the selection of XR modalities (Layer 2); these modalities can activate specific pedagogical mechanisms (Layer 3); and those mechanisms may then be evaluated and adapted through AI-driven ECD frameworks (Layer 4), together building toward a responsive educational metaverse.

3. Mechanisms of Learning: From Engagement to Stealth Assessment

3.1 Cognitive and Affective Drivers

The transition from traditional, screen-bound e-learning to the educational metaverse is often justified in terms of student engagement, but the mechanisms involved should be distinguished carefully between theoretical expectation and reported empirical tendency. Drawing on the Technology Acceptance Model (TAM) and related acceptance research, studies in this area suggest that hedonic motivation, perceived usefulness, and psychological safety are important correlates of willingness to engage with metaverse platforms [4–6]. When the learning environment incorporates spatial exploration, immediate feedback, and narrative contextualization, it may help support a state of flow. Across the current literature, however, these mechanisms are better understood as promising design rationales and reported tendencies than as uniformly demonstrated outcomes across learner groups and settings.

Furthermore, avatar embodiment may introduce the psychological dynamic often described as the "Proteus effect," in which users' behavior and self-perception can shift in relation to characteristics of their digital representation. In educational settings, this can create opportunities for strategic demographic obscuration and identity experimentation. By allowing students to alter, abstract, or neutralize their digital representations, metaverse platforms may reduce the salience of some real-world visual cues and social hierarchies. Such design choices have been associated with greater willingness to participate, experiment, and speak up, especially among learners who might otherwise experience social anxiety or stereotype threat in face-to-face settings. However, these effects are likely to be context-dependent and should be described as plausible or reported tendencies rather than as robust causal outcomes for all marginalized or introverted learners.

Beyond participation, avatar customization may strengthen psychological ownership of the learning space and reduce some forms of performance anxiety. Such affordances can encourage experimentation and low-stakes rehearsal, especially when failure can be treated as part of iterative learning rather than public embarrassment. Even here, the likely benefits are contextual rather than universal, and they should not obscure parallel concerns about identity management, accountability, and excessive gamification [7].

3.2 Assessment Mechanisms: Evidence-Centered Design (ECD) and Stealth Assessment

Traditional summative assessments can interrupt immersion by extracting learners from a simulation into a separate testing moment. One reason stealth assessment receives so much attention in educational metaverse research is that it seeks to embed evidentiary traces within the learning activity itself rather than requiring a complete break from the task [8].

Stealth assessment typically operates through Evidence-Centered Design (ECD), which links observable behavior to inferred competencies or states [8,9]. In metaverse settings, this can include interaction logs, navigation behavior, task completion patterns, dialogue traces, and, in some designs, biometric telemetry. The conceptual attraction is clear: assessment becomes continuous, contextual, and potentially adaptive. The methodological and ethical implications, however, depend heavily on what is measured, how valid those measures are, and how the resulting inferences are governed.

• Gaze Tracking and Fixation Duration: These data can indicate where visual attention is directed within a task, although attention should not be equated too quickly with deep understanding.

• Pupillary Dilation and Micro-expressions: Such signals are sometimes used as possible indicators of cognitive load or affect, but their interpretation is probabilistic and sensitive to context, lighting, and sensor quality.

• Kinematic Movement and Spatial Navigation Latency: Movement traces, hesitation patterns, and action sequences may provide evidence about problem-solving strategy, confidence, or motor proficiency in simulation-based tasks.

• Conversational Sentiment and Audio Telemetry: When learners interact with human peers or AI agents, dialogue features such as response latency, lexical choice, and help-seeking behavior can be used as additional contextual evidence.

In proposed adaptive architectures, these multi-modal traces can be processed through machine-learning models—including Bayesian student models—to update probabilistic estimates of current mastery and trigger hints or scaffold adjustments in real time [10]. This remains better understood as an emerging design direction than as a standardized capability across deployed educational metaverse systems.

4. Regional Implementations: Top-Down Strategic Deployment in China

Theoretical models become more informative when observed in institutional deployment contexts. In this review, China is treated as an illustrative leading case of state-coordinated educational metaverse development rather than as a statistically representative case of global practice. This case logic is useful because education policy, industrial strategy, infrastructure planning, and workforce development are tightly coupled in the Chinese context [11].

4.1 National Strategic Logic: The MIIT 2023-2025 Action Plan

In late 2023, China’s Ministry of Industry and Information Technology (MIIT), together with other ministries, released the Three-Year Action Plan for Innovative Development of the Metaverse Industry (2023–2025) [11]. The document elevated the metaverse from a speculative technology discourse to a coordinated policy agenda linked to industrial transformation, public-service innovation, and talent development.

The policy logic of the plan is infrastructural: the metaverse is framed not only as software but as a resource-intensive spatial internet requiring coordinated cloud, networking, computing, standards, and governance capacity. This includes 5G/6G-related connectivity, cloud-computing expansion, and interoperability frameworks intended to support persistent, synchronized virtual environments.

Crucially, the plan links immersive technologies to the “real economy,” emphasizing industrial, public-service, and training applications rather than entertainment-first consumer ecosystems. From an educational perspective, this matters because universities and vocational institutions are positioned as talent pipelines for industrial upgrading and digital-infrastructure development [11].

4.2 Higher Education: The "Metaverse Engineering" Paradigm

Section 4.2 uses the term “metaverse engineering” in a socio-technical rather than purely promotional sense. Recent mapping work defines metaverse engineering as the coordinated design of platforms, software processes, interaction systems, governance arrangements, and human roles required to build and sustain metaverse environments [12]. This broader framing helps interpret university initiatives in China more critically: some represent substantive interdisciplinary curriculum building, while others are better understood as early branding, experimentation, or institutional signaling.

- Nanjing University of Information Science & Technology (NUIST): In a widely discussed institutional move, NUIST renamed and reorganized part of its academic structure around “metaverse engineering” in 2022 [13]. The initiative is commonly described through the “BIG ANT” configuration (Blockchain, Interactivity, Game, AI, Network, IoT). In analytical terms, the case is valuable less as proof of a settled discipline than as evidence of how Chinese universities are packaging immersive infrastructure, AI, and networking into a new interdisciplinary training agenda.
- More broadly, metaverse-engineering discourse in higher education points to a shift away from treating immersive technologies as isolated teaching tools and toward viewing them as part of a larger pipeline of software development, content production, governance design, and infrastructure management [12]. This perspective also makes it easier to distinguish between empirically validated classroom outcomes and broader institutional ambitions.
- Decentralized Academic Silos: A related institutional tendency is the development of cross-campus virtual teaching spaces, shared simulation resources, and remote laboratory access. These arrangements may widen access to expertise, but their educational value remains conditional on instructional quality, interoperability, and local infrastructure rather than on the existence of a “metaverse campus” label alone.

4.3 Vocational and Technical Training: Bridging the Skills Gap

Some of the clearer empirical signals in the educational metaverse literature currently come from China’s vocational and technical sector, where embodied practice and simulation are especially relevant.

One example is the study by Wang and Siripala at Fuzhou Software Vocational and Technical College, which examined first-year digital-media-art-design students in an audio-visual language course [14]. The article reports five 45-minute sessions, pre- and post-tests, questionnaires, and 54 valid student responses. Reported post-intervention ratings were generally high (for example, means up to 4.07/5 across selected skill dimensions), and the authors also reported higher post-test than pre-test scores. Because this remains a single-institution study with limited external validation, it is best treated as preliminary empirical evidence rather than field-wide proof.

Even so, the case helps clarify why vocational education has become such a prominent application domain: many technical skills depend on repeated embodied practice, spatial troubleshooting, and low-risk rehearsal. In these contexts, immersive simulation may be especially useful because it can expand practice opportunities without the full cost, danger, or material waste of physical training environments.

4.4 Empirical Case Studies: MIET and the "Edu-Comos" Ecosystem

The operationalization of these concepts can be illustrated through selected deployments of Metaverse Intelligent Education Technology (MIET). The cases discussed below are illustrative rather than exhaustive and were chosen because they make visible different theory-to-practice pathways across higher education, language learning, teacher development, and cultural-heritage education in China. Several examples are drawn from the author's institutional context at Sanda University; they are included transparently as practice-based cases, not as neutral coverage of the entire field. **Table 1** summarizes the source basis, evidentiary status, and interpretive role of each case.

Because several of the cases in this subsection derive from institutional reports, conference papers, prototype descriptions, or author-involved practice contexts, they should be read as early-stage implementation evidence with uneven methodological depth. Limited sample reporting, context-specific deployment conditions, and restricted external validation mean that these cases are more useful for theory-to-practice illustration than for strong generalizable claims about efficacy. For readability, **Table 1** serves as a quick comparison of sector, source basis, and interpretive weight across the cases discussed below.

• Digital Twins and Immersive Learning Environments: To operationalize constructivist spatial learning, the "Edu-Comos" platform functions as a comprehensive digital twin campus [15]. This environment hosts applications ranging from interactive 3D mathematics visualization to hazard-free physics and chemistry laboratories, making it useful here primarily as an institutional design illustration rather than as a controlled outcome study.

• AI-Empowered Virtual Tutoring: AI-supported metaverse simulations have been reported as promising tools for language learning and scientific experimentation. For instance, the VIRTUSIM prototype uses software agents informed by natural-language processing and reinforcement learning to act as virtual tutors [16]. In the reported evaluation, the adaptive environment was associated with learning outcomes more than 20% higher than those obtained under a conventional teaching condition, but the available methodological detail is not sufficient here to support a stronger causal interpretation. The case is therefore best read as evidence of implementation potential for personalized scaffolding.

• Virtual Reality in Teacher Training: VR-based distance-training systems have also been deployed for in-service teachers, especially in relation to rural–urban capacity gaps [17]. The reported platform metrics—more than 1,100 concurrent users and a response time of 3.81 seconds—appear to describe a system benchmark or live implementation test rather than a controlled pedagogical experiment. Read in that light, the case is valuable because it demonstrates technical feasibility for large-scale professional development while leaving questions about learning impact open.

• Cultural Preservation via Tactile Feedback: One of the more sophisticated MIET applications links cultural-heritage education with spatial computing. The "VR Calligraphy Paradise" project combines mathematical models of brush dynamics with haptic gloves and force-feedback devices to approximate the sensorimotor dimensions of Chinese calligraphy practice [18,19]. Reported evaluations suggest gains in learners' motivation, interest, and cultural engagement, while also enabling stroke-level interaction tracking. These findings are better described as consistent with embodied-cognition-informed design than as direct validation of the theory itself.

Table 1 — Selected China-Focused Illustrative Cases in Educational Metaverse Implementation

Selected cases used for narrative synthesis; the additional evidence column clarifies source basis and interpretive weight rather than implying equivalent evidentiary status across cases.

Institution / Initiative	Sector	Application & Technology	Source Basis / Evidence Type	Interpretive Use in This Review
National MIIT Action Plan	Policy / System level	Three-Year Action Plan for Innovative Development of the Metaverse Industry (2023–2025)	Official policy document [11]	Frames national strategy, talent-development priorities, and infrastructure logic
NUIST — Metaverse Engineering	Higher Ed	Interdisciplinary metaverse-engineering curriculum framed through BIG ANT	Interview/publicity source + institutional announcement [13]	Illustrates institutional restructuring and disciplinary branding; treated as illustrative, not outcome evidence
Fuzhou Software Vocational and Technical College	Vocational / Technical	Metaverse-supported digital-media-art-design training with pre/post tests and questionnaires	Peer-reviewed single-institution empirical study [14]	Provides preliminary empirical evidence from a concrete vocational-learning context
Sanda University — Edu-Comos Platform	MIET / Higher Ed	Digital twin campus with 3D maths visualisation and virtual science laboratories	Scholarly book chapter / institutional implementation description [15]	Used as a design and deployment illustration rather than a controlled efficacy study
Sanda University — VIRTUSIM AI Tutoring	MIET / Language & Science	NLP + reinforcement-learning virtual tutors adapting to learner state	Conference paper / prototype report [16]	Suggests implementation potential for personalised scaffolding; causal claims treated cautiously
Sanda University — VR Teacher Training	MIET / Professional Development	VR distance training for in-service teachers with large-scale concurrent access	Conference paper / implementation benchmark [17]	Shows technical feasibility for scale; pedagogical impact remains context-bound
Sanda University — VR Calligraphy Paradise	MIET / Cultural Heritage	Haptic gloves, force feedback, and mathematical brush-dynamics modelling for calligraphy	Conference papers / prototype evaluations [18,19]	Illustrates embodied and cultural-heritage applications consistent with, but not proving, embodied-cognition claims

Note: MIET = Metaverse Intelligent Education Technology. Cases are purposively selected illustrative examples rather than an exhaustive dataset. The "Source Basis / Evidence Type" column is intended to differentiate policy documents, interviews/publicity materials, peer-reviewed empirical studies, and institutional implementation reports, thereby making evidentiary weight more explicit within the narrative synthesis.

5. Deployment Challenges: The Paradox of Data and Access

Despite profound pedagogical potential, systemic adoption is constrained by several highly critical logical and socio-economic paradoxes.

5.1 The Paradox of Data Governance and Privacy

The most significant operational challenge facing the educational metaverse is the inherent paradox between effective stealth assessment and user privacy. As established, the pedagogical efficacy of advanced metaverse learning relies entirely on continuous biometric telemetry. To dynamically adjust cognitive load, the system must harvest granular, real-time data, including high-frequency gaze tracking, full-body kinematic movement, pupil dilation, micro-expressions, and vocal stress analysis. However, this represents an unprecedented and highly invasive collection of sensitive, personally identifiable biological data. Unlike traditional test scores, this telemetry can inadvertently reveal profound psychological and medical insights, such as attention deficit markers, emotional volatility, or underlying neurological conditions, making its protection a critical ethical imperative.

Within the Chinese context, navigating this paradox is particularly complex. The MIIT Action Plan emphasizes data circulation, infrastructure building, and interoperable digital ecosystems [11], while privacy-centered governance frameworks require strict attention to data minimization, consent, and lawful processing. Recent work on ethical guidelines for emotion recognition in educational metaverse settings further underscores the need for purpose limitation, transparency, human oversight, and special protections when affective or biometric data are involved [20].

5.2 The Digital Divide and Cognitive Infrastructure

If immersive spatial computing becomes the baseline standard for educational quality, lack of access transforms from a mere technological inconvenience into a profound structural barrier to human capital development. As metaverse-trained students enter the workforce with vastly superior spatial reasoning and practical simulation experience, those reliant on legacy 2D education will find themselves systematically marginalized.

- **Economic and Infrastructural Barriers:** The financial threshold for metaverse integration extends far beyond the initial purchase of high-fidelity MR headsets. It necessitates continuous software licensing, rigorous hardware maintenance, and the deployment of local edge-computing servers coupled with low-latency 5G/6G networks. For rural or economically disadvantaged provinces, these infrastructure costs are prohibitively high. Without aggressive, targeted state subsidies, the metaverse threatens to rapidly exacerbate existing urban-rural educational inequalities, creating a two-tiered system where only elite, urban institutions can afford experiential learning spaces.
- **Physiological Barriers:** Prolonged exposure to synthetic environments introduces significant physiological risks, most notably “cybersickness”—a form of disorientation associated with mismatches between visual motion and vestibular feedback. In addition, when virtual environments are poorly optimized or overly complex, the novelty of navigation and interface management can consume cognitive resources that should be directed toward the learning task. Poor frame rates, unstable interaction design, and demanding locomotion mechanics therefore remain practical design constraints rather than minor usability details.
- **Psychological and Social Barriers:** The hyper-stimulating nature of fully immersive environments presents profound psychological risks. The phenomenon

known as the "Proteus effect"—where user behavior is unconsciously altered by their digital avatar's characteristics—can lead to identity confusion, especially among younger learners whose physical identities are still forming. Additionally, the gamified, highly engaging nature of the educational metaverse significantly increases the risk of technological addiction and escapism, potentially rendering the analog classroom—and real-world social interactions—under-stimulating and mundane by comparison.

To mitigate these multifaceted risks, educational institutions and policymakers must move beyond mere hardware procurement. They must concurrently deploy digital-wellness guidelines, strict immersion time limits, physical-world integration activities, and psychological support systems. In this sense, the educational metaverse should be governed as a socio-technical system rather than as a headset rollout. Formula 1 and the accompanying variable-definition table summarize this tension in heuristic form.

Formula 1 — Conceptual Heuristic for Educational Metaverse Effectiveness (EME) in the AI Era

An illustrative synthesis of cognitive, technological, institutional, and ethical dimensions

$$EME(t) = \alpha \bullet (EC \times XRf) + \beta \bullet (AIa \times ECDd) + \gamma \bullet (PA \times IC) - \Delta \bullet (DDg + Pr)$$

Conceptual note: Formula 1 is presented as a heuristic device for organizing the discussion, not as a validated predictive model or directly operationalized index. The variables below should therefore be read as candidate constructs that would require explicit observable proxies—for example, presence or engagement scales, telemetry richness, infrastructure-readiness indicators, and privacy-burden measures—before empirical measurement.

Illustrative Variable Definitions

$EME(t)$	Educational Metaverse Effectiveness at time t — an illustrative composite construct used for conceptual synthesis rather than direct empirical measurement in its current form
α	Weighting coefficient for theoretical-cognitive dimension ($0 < \alpha < 1$)
EC	Embodied Cognition engagement score — measures motor-cortex co-activation during learning (0–1 scale)
XRf	Extended Reality Fidelity — immersion depth of the XR modality used (AR < Lifelogging < Mirror World < VR)
β	Weighting coefficient for AI-adaptive dimension ($0 < \beta < 1$)
AIa	AI Adaptation coefficient — real-time personalisation rate of the LLM/Bayesian student model (0–1 scale)
$ECDd$	ECD Depth — richness of biometric telemetry streams feeding Evidence-Centred Design (0–1 scale)
γ	Weighting coefficient for policy-infrastructure dimension ($0 < \gamma < 1$)
PA	Policy Alignment index — degree of state/institutional mandate coherence (e.g., MIIT Action Plan conformance)
IC	Infrastructure Capacity — network bandwidth × compute × XR hardware penetration rate
Δ	Attenuation factor ($\Delta \geq 0$) — combined drag of equity and governance constraints

DDg	Digital Divide gap — socioeconomic disparity in hardware access across learner populations
Pr	Privacy Risk burden — compliance cost and ethical friction of biometric data governance (e.g., PIPL)

Design proposition: For a given learning objective, *EME(t)* would be expected to improve when weighting coefficients are specified transparently, when XR fidelity and ECD depth are matched to pedagogical purpose, when adaptive support is helpful rather than excessive, and when the combined burden of digital divide and privacy risk is reduced. This is intended as a hypothetical design proposition rather than as a formally derived optimization theorem.

Interpretation: The heuristic captures a central tension in discussions of the AI-era educational metaverse. The additive terms represent potential value-creating forces—cognitive embodiment, AI-supported personalization, and policy-backed infrastructure—while the subtractive term encodes the governance paradox that richer telemetry can also intensify privacy risk and exclusion. Approaches such as federated learning, differential privacy, and data minimization are therefore best understood here as governance-design options that future systems may operationalize rather than as solutions already demonstrated across the evidence reviewed.

6. Conclusion

The educational metaverse can be understood as a possible shift from predominantly screen-based learning toward more spatial, embodied, and simulation-rich forms of educational interaction. Across the literature reviewed here, constructivist pedagogy, embodied cognition, and stealth assessment provide a useful conceptual vocabulary for explaining why immersive environments may support conceptual understanding, practice, and engagement [1–3,8–10]. At the same time, the strength of the evidence varies considerably across studies and settings. The main contribution of this article is therefore to synthesize the field's dominant theoretical logics and connect them with emerging implementation examples, rather than to claim that the educational value of the metaverse has already been definitively established.

As this narrative review suggests, China provides an especially informative case for examining the systemic, state-led integration of educational metaverse technologies. The cases discussed here indicate how immersive initiatives can be aligned with wider goals such as industrial upgrading, teacher development, and cultural preservation. At the same time, these examples remain heterogeneous in source type, methodological rigor, and evidentiary strength. The higher education, vocational, tutoring, teacher-training, and calligraphy cases therefore provide promising but context-bound evidence of what MIET systems may enable, rather than definitive proof of uniform improvement across settings [11–19].

However, realizing this potential is neither guaranteed nor frictionless. It requires resolving major ethical and socio-economic tensions. Chief among these is the conflict between adaptive, telemetry-rich assessment and privacy-centered governance, especially when gaze, affective, or behavioral data are collected continuously [8–10,20]. At the same time, the infrastructural demands of high-bandwidth networks, edge computing, and XR hardware risk deepening educational inequality if access remains concentrated in better-resourced institutions. The longer-term value of the educational metaverse will therefore depend less on technological novelty alone than on whether institutions can balance pedagogical ambition with proportionate governance, accessibility, and evidence building.

Future empirical work should move beyond proof-of-concept reporting toward more comparable and methodologically transparent designs. Priorities include multi-site or cross-national studies, clearer operationalization of constructs such as presence, engagement, learning transfer, and telemetry-based inference, more explicit separation between system-benchmark evidence and pedagogical outcome evidence, and longer-term evaluation of privacy, equity, and governance trade-offs in XR-supported learning environments.

Author contribution statement: The sole author YDX was responsible for the conceptualization, methodology, formal analysis, investigation, writing – original draft preparation, writing – review and editing, visualization, and project administration. All the work was performed by the author alone.

Funding: This research was funded by the Shanghai Municipal Education Commission Educational Science Planning, grant number C2023035, for the project "Theoretical Construction and Exploratory Application of the Educational Metaverse".

Conflict of Interest: The author declares no conflict of interest.

References

1. Kye B, Han N, Kim E, et al. Educational applications of metaverse: Possibilities and limitations. Journal of Educational Evaluation in Health Professions. 2021, 18: 32. https://doi.org/10.3352/jeehp.2021.18.32
2. Zhang X, Chen Y, Hu L, et al. The metaverse in education: Definition, framework, features, potential applications, challenges, and future research topics. Frontiers in Psychology. 2022, 13: 1016300. https://doi.org/10.3389/fpsyg.2022.1016300
3. Sin ZPT, Jia Y, Wu ACH, et al. Toward an Edu-Metaverse of Knowledge: Immersive exploration of university courses. IEEE Transactions on Learning Technologies. 2023, 16(6): 1096–1110. https://doi.org/10.1109/TLT.2023.3290814
4. Alkhwaldi AF. Investigating the social sustainability of immersive virtual technologies in higher educational institutions: Students' perceptions toward metaverse technology. Sustainability. 2024, 16(2): 934. https://doi.org/10.3390/su16020934
5. Liu Y, Huang J, Saleh S, et al. A review of the application of the metaverse in education. Educational Administration: Theory and Practice. 2024, 30(8): 84–97. https://doi.org/10.53555/kuey.v30i8.6986
6. Wu R, Yu Z. Investigating users' acceptance of the metaverse with an extended technology acceptance model. International Journal of Human–Computer Interaction. 2024, 40(19): 5810–5826. https://doi.org/10.1080/10447318.2023.2241295
7. Turkay S, Kinzer CK. The effects of avatar-based customization on player identification. International Journal of Gaming and Computer-Mediated Simulations. 2014, 6(1): 1–26. https://doi.org/10.4018/IJGCMS.2014010101
8. Shute VJ, Ventura M. Stealth Assessment: Measuring and Supporting Learning in Video Games. MIT Press; 2013. https://doi.org/10.7551/mitpress/9589.001.0001
9. Shute VJ, Wang L, Greiff S, et al. Measuring problem solving skills via stealth assessment in an engaging video game. Computers in Human Behavior. 2016, 63: 106–117. https://doi.org/10.1016/j.chb.2016.05.047
10. Hadyaoui A, Cheniti-Belcadhi L. AI-driven adaptive stealth assessment for socially regulated learning in collaborative environments. Journal of Research on Technology in Education. 2026. https://doi.org/10.1080/15391523.2025.2586501
11. Ministry of Industry and Information Technology of the People's Republic of China, Ministry of Education of the People's Republic of China, Ministry of Culture and Tourism of the People's Republic of China, et al. Three-Year Action Plan for Innovative Development of the Metaverse Industry (2023–2025). 2023. Available online: https://gxt.fujian.gov.cn/zwgk/ghjh/202309/P020230908621291575052.pdf (accessed on 31 March 2026).
12. Di Dario D, Palomba F, Gravino C, et al. Another brick in the wall: A systematic mapping study toward defining metaverse engineering through socio-technical issues. ACM Computing Surveys. 2026. https://doi.org/10.1145/3785659
13. Pan Z. How to build the first metaverse department in Chinese universities? Dialogue with Prof. Zhigeng Pan, editor-in-chief of Metaverse. 2022. Available online: https://aber.apacsci.com/index.php/met/announcement/view/172 (accessed on 31 March 2026).

14. Wang Y, Siripala W. Metaverse enhancing vocational skills of digital media education in Chinese higher vocational institutions. Journal of Education and Learning. 2025, 14(5): 239–249. https://doi.org/10.5539/jel.v14n5p239
15. Yu D. Designing effective learning environments in the educational metaverse: The role of augmented and virtual reality. In: Geroimenko V, editor. Augmented and Virtual Reality in the Metaverse. Springer; 2024. pp. 81–100. https://doi.org/10.1007/978-3-031-57746-8_5
16. Yu D. AI-empowered metaverse learning simulation technology application. In: 2023 International Conference on Intelligent Metaverse Technologies & Applications (iMETA). IEEE; 2023. pp. 1–6. https://doi.org/10.1109/iMETA59369.2023.10294830
17. Yu D. Practical exploration of educational metaverse—A teacher training system based on virtual reality. In: 2022 10th International Conference on Orange Technology (ICOT). IEEE; 2022. pp. 1–4. https://doi.org/10.1109/ICOT56925.2022.10008186
18. Ma N, Yu D. How do VR applications address the challenges of conveying the tactile feedback crucial to traditional calligraphy practice? In: Proceedings of the 16th International Conference on Computer Supported Education, Volume 1: ERSeGEL. SciTePress; 2024. pp. 744–750. https://doi.org/10.5220/0012753600003693
19. Yu D, Ma N. Enhancing Chinese character education and calligraphy through metaverse intelligent technology: Integration and impact. In: Hong W, Kanaparan G, editors. Computer Science and Education. Educational Digitalization. CCIS 2025. Springer; 2024. pp. 181–193. https://doi.org/10.1007/978-981-97-0737-9_18
20. Di Dario D, Pentangelo V, Colella MI, et al. Collecting and implementing ethical guidelines for emotion recognition in an educational metaverse. In: Adjunct Proceedings of the 32nd ACM Conference on User Modeling, Adaptation and Personalization (UMAP Adjunct '24). ACM; 2024. https://doi.org/10.1145/3631700.3665232