

---

# Construction of an Embodied-Cognition-Based Classroom Teaching Evaluation Model Using Multimodal Perception

Zhifeng Wu<sup>1\*</sup>

<sup>1</sup> Guangdong Polytechnic Normal University, Guangzhou, 510400, China

\* 1005520633@qq.com

<https://doi.org/10.70695/IAAI202602A8>

---

## Abstract

Traditional classroom teaching evaluation is deeply influenced by disembodied rationalism, spatiotemporal fragmentation, and technological alienation. Drawing upon embodied cognition theory and digital intelligence technologies, this study constructs an "embodied cognition" classroom teaching behavior evaluation model supported by multimodal perception. The model evaluates the practice of teachers and students across hybrid physical-virtual scenarios via five core dimensions: situational reconstruction, dynamic generation, bidirectional adaptation, verbal embodiment, and fusion of horizons. By orchestrating multi-source dynamic signals, the model achieves a paradigm shift in classroom evaluation from "post-hoc lag" to "dynamic symbiosis", while rationally empowering human teachers' "practical knowledge" and "educational empathy". This research aims to solidify a human-centered evaluative foundation, steering education toward its authentic mission of cultivating well-rounded individuals under the integration of science and education.

**Keywords** Embodied Cognition; Multimodal Data Capture; Classroom Teaching Evaluation; Digital Intelligence Transformation

---

## 1 Mirror and Fracture: Manifestations, Limitations, and Theoretical Reflections on Traditional "Disembodied" Classroom Teaching Evaluation

### 1.1 Academic Scrutiny of Cognitive Rationalism: The "Disembodied" Ontological Representation of Traditional Classroom Evaluation

The intellectual genealogy of traditional classroom evaluation heavily inherits Cartesian "mind-body dualism", which mechanically reduces complex teaching activities to pure processes of information processing and symbolic computation. Operating under the computational metaphor of "the brain as a computer", this perspective demotes the physical body to an isolated tool of execution or a biological container completely detached from emotional and cognitive states. Consequently, teaching evaluation degenerates into a rigid "pre-set inspection" focused solely on the retention of explicit knowledge. Traditional methods of classroom teaching behavior analysis rely predominantly on manual coding and time-interval sampling to gather behavioral data. While this approach historically contributed to the standardization of classroom quality, manual sampling possesses inherent methodological flaws: small sample sizes, labor-intensive workflows, high error rates, and pronounced subjective biases. These deficiencies obscure genuine instructional issues and underlying pedagogical laws [1]. Furthermore, such methods strip teaching of its holistic nature as an embodied cognitive practice. By disregarding the tacit knowledge continuously co-created by teachers and students through physical interactions within specific spatial and cultural fields, traditional evaluation remains frozen at the level of superficial symbolic measurement. It fails to capture the rich cognitive experiences of students as fully integrated cognitive agents, steering classroom evaluation away from complex vital realities and toward a highly mechanized, reductionist path.

### 1.2 Empirical Intuition and Subjective Obscuration: The Methodological Limitations of Manual Qualitative Evaluation

In terms of evaluative instruments and methodologies, traditional classroom evaluation relies heavily on evaluators' empirical intuition, trapping the process in a closed loop of subjectivity. Because

traditional cognitivism upholds a strict mind-body separation, cognitive processes are viewed as abstract computations performed independently of bodily states [2]. When classroom evaluation depends entirely on the naked eye and post-hoc qualitative inferences, it inevitably falls into the trap of subjective bias. This observer-centric, retrospective qualitative paradigm is highly vulnerable to cognitive distortions such as the "halo effect" and the "recency effect". Evaluators frequently utilize isolated, static "behavioral correlation nodes"—such as the frequency of a teacher's questions, specific blackboard configurations, or student hand-raising rates—to make broad empirical inferences regarding complex student emotions and classroom psychological safety. Such sensorially constrained evaluations not only mask fleeting yet critical details of embodied interaction—such as cognitive resolution at the intersection of teacher-student perception or forward trunk leaning indicating deep cognitive engagement—but also introduce a high degree of intentionality and non-replicability. Consequently, they fail to provide the precise diagnostic feedback necessary for optimizing instructional designs.

### **1.3 Spatial-Temporal Fragmentation and Generative Rupture: The Discretization of Dynamic Teaching Flows by Quantitative Evaluation**

From a cognitive standpoint, traditional evaluations of classroom teaching exhibit a severe manifestation of "spatial-temporal fragmentation", which precipitates a profound rupture in generative learning. Traditional evaluation frameworks predominantly adopt a "sliced" or "summative" logic, isolating instructional behaviors within highly restrictive temporal nodes or rigid spatial dimensions. This logic violently dissects the continuous, fluid, and symmetrical "life flow" of the classroom into fragmented, commodified "instructional interruptions [3]". In stark contrast, embodied cognition theory asserts that the physical body actively shapes cognition through rich interactions involving emotion, situated context, and symbolic environments. Within this framework, motor perception provides the foundational mechanism for bodily presence, experiential focus reinforces emotionally driven care, and field configuration organizes the generation of interwoven virtual-physical contexts. By fracturing the spatial-temporal continuum, traditional evaluations strip students of the physical context necessary to perceive their own feelings and emotions, rendering it impossible to trace cognitive trajectories within the dynamic interactions among bodies, environments, and tools. When evaluation remains obsessed with the "completion rate of pre-set instructional objectives", spontaneous creative generation arising from emotional convergence and flashes of "insight" triggered by cognitive conflict are systematically filtered out as mere "technical noise" because they cannot be accommodated by standardized evaluation metrics [4].

## **2 Shift and Reshaping: The Theoretical Architecture of the "Embodied Cognition" Classroom Teaching Evaluation Model Based on Multimodal Perception**

### **2.1 Physical Fields and Embodied Coupling: The "Situational Reconstruction" Dimension of Environment-Body Interaction**

Viewed through the lens of embodied cognition, the classroom is never a mere physical container housing instructional tools, but a dynamic field intrinsically bound to the vital existence of teachers and students. Traditional evaluation treats the environment as a passive background; however, the multimodal cognitive model prioritizes "embodied coupling". Rooted in the affordance theory of ecological psychology, this dimension posits that cognition emerges from the harmonious evolution of mind, body, and physical surroundings. Classroom variables—such as desk arrangements, multimedia layouts, and behavioral trajectories—exert real-time structural influences on the somatic and cognitive pathways of both teachers and students [5]. By capturing spatial topological maps, the frequency of instructional tool utilization, and the self-adaptive adjustments of the classroom microclimate, this model comprehensively demonstrates how a physical field transforms into a supportive ecology that fosters deep learning. The integration of digital intelligence allows classroom evaluation to dynamically track student developmental trajectories while executing large-scale data storage and analytics, enabling qualitative behavioral descriptions that were previously unattainable. This dimension successfully shifts teaching evaluation from isolated individual measurements to systematic interventions targeting the interactive quality of the interconnected "human-environment" system.

## **2.2 Bodily Metaphors and Cognitive Emergence: The "Dynamic Generation" Dimension of Multimodal Behavioral Trajectories**

The physical body itself acts as the primary instigator of cognitive processes. The dimension of "dynamic generation" centers on the concepts of bodily metaphors and cognitive emergence, purposefully breaking the traditional evaluation paradigm's obsession with static, purely mental outputs. Embodied cognition theory indicates that the comprehension of abstract concepts relies heavily on the sensorimotor system; therefore, the micro-expressions, gestures, bodily postures, and somatic tensions of teachers and students are not secondary byproducts of cognition, but rather the embodied manifestations and real-time regulators of thinking processes. For instance, the posture of courtesy and the embodied metaphorical transformation of logical concepts can be precisely mapped; a distinct forward lean of the torso reflects specific cognitive engagement models. By utilizing multimodal sensors to capture these subtle, dynamic bodily data, the model converts the historically unmeasurable "implicit cognitive flow" into quantifiable, analyzable embodied dynamic characteristics. This provides educators with actionable evidence to enrich instructional content, construct comprehensive knowledge taxonomies, optimize instructional organization, and map out precise implementation pathways — ultimately generating evidence-based feedback to elevate teaching standards and enhance the cultivation of innovative talents [6].

## **2.3 Intersubjectivity and Emotional Resonance: The "Bidirectional Adaptation" Dimension of Physiological and Micro-expression Fusion**

Teaching is fundamentally an intersubjective process realized through the physical co-presence of students and teachers within a shared field. The dimension of bidirectional adaptation focuses on emotional resonance within the classroom ecology and the affective dynamics defined by social neuroscience. In an embodied instructional field, teachers and students establish cross-subjective emotional bonds mediated by the mirror neuron system. Under this paradigm, the evaluative collective expands to include students, peer teachers, instructional supervisors, and the instructors themselves, generating a comprehensive multi-perspective matrix [7]. Students, as the direct recipients and active participants of instruction, serve as an invaluable source of evaluative data, facilitating a holistic understanding of the classroom climate. The evaluation model rejects unidirectional emotional output from the teacher, opting instead for a "dynamic somatic adaptation" that tracks bodily language and physiological arousal across all participants. Through the synchronized analysis of facial micro-expressions and physiological signals captured via multimodal streams, the system accurately identifies emotional frequencies and shared resonance zones. This technological breakthrough converts subjective emotional hypotheses into verifiable data patterns, providing an empirical pathway to assess psychological safety, evaluate empathetic efficacy, and manifest humanistic care through technical execution.

## **2.4 Discourse Spectra and Interaction Depth: The "Verbal Embodiment" Dimension of Multi-Spatial-Temporal Dialogue Perception**

Speech is not an abstract symbol floating above the physical body; rather, it possesses profound physical properties that constitute "verbal embodiment". Grounded in speech act theory, this dimension acknowledges that the semantic essence of language is highly coupled with physical acoustic features, such as tonal inflections, speech rate fluctuations, and volume intensity, as well as accompanying physical postures. While current research in embodied AI regarding human somatic language primarily concentrates on facial node identification and driving technologies to decode basic expressions, this model incorporates advanced voice recognition and natural language processing (NLP) to perform deeper diagnostics. The evaluation framework analyzes not only the logical hierarchy, cognitive structures, and ladders of classroom semantics but also integrates multi-source acoustic characteristics into its assessment matrix [8]. By capturing vocal pauses during a teacher's lecture and tonal shifts during student responses, the system delineates a highly accurate discourse energy spectrum. This spectrum diagnoses whether classroom interaction remains stalled at superficial conceptual descriptions or has successfully progressed to an advanced cognitive stage characterized by deep somatic dialogue, perspective negotiation, and high-order critical thinking.

## **2.5 Intentional Focus and Co-Attention: The "Fusion of Horizons" Dimension of Multi-Source Eye-Tracking**

The allocation of visual attention represents the most direct, multimodal index of cognitive engagement. The dimension of the "fusion of horizons" relies on intentional focus and joint attention theories to construct a visual spatial topology for classroom evaluation. Embodied cognition posits that an individual's gaze vector constitutes the explicit boundary of their conscious flow and cognitive representation. In traditional disembodied evaluation paradigms, attention is often superficialized as a macro-level "head-up rate", which lacks diagnostic precision. The multimodal cognitive model leverages advanced eye-tracking algorithms to capture the precise gaze characteristics of both teachers and students. By evaluating the alignment of joint attention on specific instructional targets, gaze fixation durations, and the dynamic features of gaze switching, the system determines whether an authentic "fusion of horizons" has been achieved between subjects. This perspective emphasizes that the application of artificial intelligence in educational governance does not exist in isolation; instead, it is deeply embedded within existing educational governance structures and societal environments, shaped by an array of intersecting variables [9]. This dimension fundamentally elevates attention evaluation from a superficial check of classroom order to a precise diagnostic focus on the co-construction of cognitive states.

## **3 Field Innovation: Practical Landscapes and Application Scenarios of the "Embodied Cognition" Classroom Teaching Evaluation Model**

### **3.1 Micro-Adjustments: Real-Time Regulation Scenarios of Teacher-Student "Emotion-Cognition" Dynamic Resonance**

For a long time, traditional educational assessment has relied on post-hoc evaluations. This retrospective feedback mechanism suffers from a structural latency, making it impossible to respond immediately to student confusion or shifting classroom dynamics as they occur. In modern smart classrooms empowered by "new quality productive forces", the ubiquitous embedding of digital intelligence completely shatters this "black box". By deploying non-intrusive sensing arrays, the system continuously monitors the classroom environment and translates non-verbal symbols into real-time, visualized emotional and attentional indicators. Through the non-invasive tracking of student gaze vectors, facial expressions, and interaction frequencies, the system generates instantaneous diagnostic reports. Instructors can leverage these insights to implement agile, dynamic pedagogical adjustments and flexible instructional interventions. This highly responsive regulation mechanism resolves the cognitive fatigue caused by mismatched teaching paces, eliminates emotional-cognitive disconnects, and transforms the classroom into a premier physical site where all educational elements are optimally configured. Consequently, students shift from passive consumers of knowledge to autonomous agents of learning, provided with rich sensory materials that deepen their cognitive processing. This directly aligns with the philosophy of embodied instruction, which advocates transforming passive learning into active practice, abandoning rote-learning paradigms, fostering intrinsic motivation, and promoting the seamless integration of bodily experience and cognitive growth [10].

### **3.2 Spatial Turn: Efficacy Manifestation Scenarios of "Multidimensional Embodied Operations" in Inquiry-Based Learning**

In traditional inquiry-based learning settings, the lack of pervasive behavioral datasets often causes educational evaluation to drift into vague qualitative descriptions, leaving inquiry-based courses vulnerable to accusations of unsubstantiated learning efficacy. New instructional spaces driven by new quality productive forces resolve this dilemma by embedding intelligent sensor networks and edge computing architectures, achieving a holographic deconstruction of student spatial behaviors. This scenario utilizes indoor motion capture and spatial layout tracking technologies to collect real-time data regarding students' physical proximities and exchange modalities during collaborative tasks, precisely extracting metrics such as tool manipulation frequency and physical interaction complexity. These heterogeneous behavioral elements are mathematically compiled into quantifiable embodied cognitive streams, providing a rigorous representation of implicit engagement levels and current knowledge depth. Based on these data streams, the system iteratively optimizes spatial designs and dynamically

reorganizes resources, upgrading the physical field from a passive instructional container to a highly flexible, borderless intelligent space that empowers the perfect unity of knowing and doing.

### **3.3 Co-Symbiosis: Collective Intelligence Emergence Evaluation Scenarios Based on "Joint Action" in Group Collaboration**

Traditional evaluations of collaborative learning heavily favor individual performances, frequently overemphasizing solo verbal contributions while ignoring critical collaborative dimensions such as non-verbal coordination and somatic tacit understanding. Furthermore, conventional formative evaluation techniques place immense pressure on instructors, requiring them to possess advanced evidence-based mindsets while simultaneously observing and recording process-based behaviors, which severely complicates practical implementation [11]. By leveraging a robust digital intelligence network, this scenario shifts the evaluative lens entirely toward group-level interactions. The system combines machine vision and semantic space analysis to monitor the group's somatic pedigree and dynamic interactive networks in real time, capturing subtle variables like gaze convergence and physiological synchronization among group members. By deeply deconstructing the joint actions and cognitive efficiencies of heterogeneous agents facing complex problem-solving tasks, the system precisely quantifies the transition of a team from mere "physical interaction" to an integrated "chemical reaction". This breakthrough dismantles isolated individual evaluation paradigms, achieving a major capacity leap from unquantifiable group dynamics to structured, measurable collective intelligence, while comprehensively mapping out the harmonious co-evolution of cross-subjective collaborations.

### **3.4 Individual Profiles: Panoramic Developmental Evaluation Scenarios of Dynamic Embodied Learning Toward the "Unity of Knowing and Doing"**

Traditional assessment mechanisms have long been crippled by an over-reliance on exam scores and summative outcomes, leaving them incapable of mapping a student's long-term capability evolution within complex research and learning fields. The profound integration of human instruction and artificial intelligence is paramount; within this space, students and teachers establish relationships with AI that transcend basic utility, forming unique, co-adaptive bonds. This scenario leverages ubiquitous digital intelligence to execute long-term, cross-field multimodal tracking of individual learners. By merging heterogeneous data sources into a unified analytics platform, the system dynamically charts a student's classroom participation, emotional variance, and focus indices, extracting and visualizing implicit cognitive cycles and workflow progressions. Through the deep integration of scientific cognition, collaborative actions, and physiological markers, the system generates multidimensional, stereoscopic panoramic profiles that accurately capture an individual's energetic transitions during knowledge construction. This panoramic developmental evaluation dismantles binary administrative metrics, providing the precise data infrastructure required to deliver authentic individualized education and pushing integrated education toward true praxis.

## **4 Transcendence and Guarding: Role Reshaping and Value Reflections of Human Teachers in the Era of Embodied Intelligent Evaluation**

### **4.1 Digital Demystification of Empirical Intuition: Rational Empowerment of Teachers' "Practical Knowledge" via Embodied Evaluation**

Through long-term professional practice, human teachers accumulate rich instructional experiences; yet, these experiences are inherently limited by cognitive biases, routine thinking, and perceptual distortions that can induce sporadic instructional errors. Evaluation tools driven by new quality productive forces perform a profound "digital demystification" of the traditional classroom through holographic multimodal technologies, utilizing objective statistical data to correct embedded human biases. The ubiquitous introduction of data intelligence streams precisely illuminates blind spots in teacher self-perception. Crucially, this structural reconfiguration does not negate traditional educational wisdom; rather, it catalyzes a paradigm shift summarized by the formula:  $\text{Experience} + \text{Data} = \text{Precision Diagnosis}$ . The strategic integration of data assets encourages teachers to transition from isolated individual awareness to a shared community consciousness, shifting from private rationality to the public rationality of a professional community [12]. This deepens collaboration

and consensus among educators, driving the iterative upgrading of their practical knowledge and achieving a profound rational empowerment of instructional reconstruction.

#### **4.2 The Humanistic Return of Technological Rationality: Ecological Nourishment of Teachers' "Educational Empathy" via Embodied Data**

As new quality productive forces integrate into the modern industrial system—driving profound structural transformations and industrial upgrades across manufacturing sectors—education remains a unique domain where human empathy must serve as the core pedagogical foundation and must never be dismantled [13]. Tech leaders have long noted that empathy is among the most difficult skills to cultivate or replicate. In the smart era, technological tools (such as AI-driven educational software and virtual/augmented reality systems) are widely deployed, making screen-based and VR/AR interactions a dominant modality. This heavy reliance on technological mediation can inadvertently constrain emotional development and empathy cultivation. However, multimodal embodied data, rather than sliding into a cold, technocentric paradigm, acts as a digital bridge to help teachers perceive the inner psychological worlds of their students by converting implicit emotional and stress states into explicit, actionable data. This holographic insight enables teachers to look past outward behavioral surfaces and see true implicit needs—accurately reading the anxiety or depression hidden behind a student's silence [14]. Here, technology's ultimate value—serving the understanding of human ontology—is fully realized. It liberates teachers from mechanical routines like attendance checking and rote information delivery, steering educational practices back to their humanistic origins. Data, in this light, is transformed into a current of systemic goodwill, ecologically nourishing the warmth of education.

#### **4.3 Pedagogical Reflection on Bodily Politics: Intellectual Evolution and Self-Iteration of Teachers Based on Embodied Visions**

Traditional instructional spaces are frequently plagued by an entrenched inertia of bodily discipline, where rigid spatial layouts and unidirectional gazes reflect deep-seated power dynamics. Within borderless intelligent networks driven by new quality productive forces, embodied evaluation compels teachers to become acutely aware of their own physical presence, prompting them to reflect on how their spatial positioning, physical intent, and vocal tones subtly shape student cognitive paths. Supported by digital intelligence, teachers learn to accurately decode the participation or alienation signaled by subtle somatic clues, resolving structural latencies in knowledge transmission and advancing classroom evaluation from a purely data-driven model to a paradigm of human-machine symbiosis. This shift directly serves the core objectives of evaluation reform and meets the fundamental needs of professional teacher development [15]. Through this systematic deconstruction under an embodied vision, teachers shed the role of authoritarian lecturers, executing an intellectual evolution and self-iteration to become facilitators of an integrated educational ecosystem.

#### **4.4 Safeguarding the Authenticity of Humanism: Guarding Against the Traps of the "Data Panopticon" and Disembodied Technological Alienation**

If digital intelligence technologies are deployed without rigorous ethical paradigms, they risk creating severe data surveillance crises, alienating the classroom into a digital "panopticon". Educators must remain highly vigilant against the mechanical dissection of human cognition by over-quantification tools, strictly guarding against capital-driven logics that reduce human nature to cold, simplistic indices [16]. It must be unequivocally asserted that technology is merely an instrument and a means to an end. Educational institutions must unswervingly adhere to ethical frameworks and privacy protection regulations, building a human-centered, human-machine collaborative, accountable, and secure intelligent educational governance ecology by enforcing technological transparency and controllability [17]. This ensures that digital data streams consistently honor the holistic nature of human life, returning faithfully to the original mission of education and the authentic value of social-pedagogical relations. The construction of an embodied classroom teaching evaluation model based on multimodal perception represents a highly focused, intelligent utilization of multi-source data under the strategic framework of "Digital China", charting a clear implementation path for intelligent agents to empower national digital development [18]. In the near future, intelligent perception arrays will seamlessly permeate physical learning spaces, achieving ubiquitous learning context awareness. The continuous refinement of this

model will solidify the core value foundation that "evaluation must serve human development". Concurrently, as a critical instructional hub empowering the actualization of new quality productive forces, embodied intelligent evaluation will break through bureaucratic fragmentations and rigid educational discipline, championing human-machine synergy to better benefit instructional design and evaluation practices [19]. By allowing digital technology to routinely serve holistic educational practices and vital human growth, this framework steers higher education toward an authentic return to humanism and science-education integration under a life-oriented, modern paradigm.

## Acknowledgement

This work was supported without any funding.

## Conflicts of Interest

The authors declare no conflicts of interest.

## References

1. Zhang, L. L., & Gu, X. Q. (2022). Analysis model and practice framework of classroom teaching behavior supported by multimodal data. *Open Education Research*, 28(6), 101-110.
2. Yang, X., & Mao, Y. H. (2025). Tech-dependence anxieties of science teachers on GenAI agents in the perspective of embodied cognition and the return of subjectivity. *Journal of Teacher Education*, 12(5), 32-40.
3. Qiu, L. J. (2026). Research on the digital transformation of classroom teaching from the perspective of embodied cognition theory. *Theory and Practice of Education*, 46(8), 57-64.
4. Li, K. Y. (2025). Weak embodiment, embodied cognition and embodied AI. *Studies in Dialectics of Nature*, 41(3), 81-87.
5. Grp, C. H., & Yu, Z. Y. (2025). Practical dilemmas and optimization paths of classroom teaching evaluation empowered by digital intelligence. *Journal of the Chinese Society of Education*, (7), 51-58.
6. Song, Y., Xu, C. L., & Mu, X. X. (2024). Research on new classroom teaching evaluation and optimization empowered by generative artificial intelligence. *Modern Educational Technology*, 34(12), 27-36.
7. Kang, L. Y., Ye, H. Z., Zhang, J., et al. (2022). Research on classroom teaching behavior evaluation based on multi-dimensional data. *Modern Educational Technology*, 32(4), 68-77.
8. Wu, M. X., & Chen, H. J. (2025). Phenomenological reflections on the interaction behavior of embodied intelligence. *Studies in Dialectics of Nature*, 41(10), 79-84.
9. Luo, Y. (2025). AI-driven educational governance transformation: From the perspective of attention allocation. *Northwest Journal of Ethnology*, (6), 124-133.
10. Zhao, X. F., & Wu, Y. X. (2025). Integration and mutual learning between multimodal pedagogy and cognio-somatic linguistics. *Foreign Languages in China*, 22(4), 4-10, 22.
11. Zhang, L. S., Feng, S., & Li, T. T. (2021). Formal modeling and intelligent computing for classroom teaching evaluation. *Modern Distance Education Research*, 33(1), 13-25.
12. Zhang, J., & Fan, G. X. (2026). Return of public rationality in teacher evaluation in the digital intelligence era. *China Examinations*, (5), 19-28.
13. Chen, N. J., Men, X. H., & Chen, X. (2025). Demand transformation and paradigm reconstruction of talent cultivation in the field of intelligent manufacturing under the perspective of embodied intelligence. *China University Teaching*, (7), 46-53.
14. Wang, Z. J., Liu, X. J., & Su, C. Y. (2025). How humans differ from AI: The cultivation path of empathy based on design thinking in intelligent society. *Chinese Journal of Distance Education*, 45(6), 20-34.
15. Wang, Z. J., Long, S., & Zhang, J. (2025). Research on the construction of hierarchical model of human-machine collaborative intelligent classroom teaching evaluation. *Distance Education Journal*, 43(5), 32-40.
16. Zhang, Q., Luo, X., & Chen, Y. J. (2025). Research on educational big data-driven student development situational awareness and early warning system. *Modern Distance Education*, (6), 53-63.
17. Hu, X. S., Fu, Q., Huang, X. L., et al. (2026). Construction of intelligence agent model driven by multimodal data. *Journal of Intelligence*, 1-9. <https://link.cnki.net/urlid/61.1167.g3.20260114.1540.002>
18. Wang, Y. Y., & Zheng, Y. H. (2022). Multimodal data fusion: The core driving force for solving key problems of intelligent education. *Modern Distance Education Research*, 34(2), 93-102.
19. Yu, S. R., & Xiong, S. S. (2024). Architecture of artificial general intelligence teacher enhanced by large models. *Open Education Research*, 30(1), 33-43.

## Biographies

1. **Zhifeng Wu** Master's student at Guangdong Polytechnic Normal University.

## 基於多模態感知的具身認知課堂教學評價模型構建

吳治鋒<sup>1</sup>

<sup>1</sup>廣東技術師範大學，廣州，中國，510400

---

**摘要：**傳統的課堂教學評估深受非具身理性主義、時空碎片化以及技術異化的影響。本研究借鑒具身認知理論與數位智慧技術，建構了一個基於多模態感知的「具身認知」課堂教學行為評估模型。該模型透過五個核心維度（情境重構、動態生成、雙向適應、言語具身化及視野融合），對混合式物理—虛擬場景中師生的教學實踐進行評估。透過整合多源動態訊號，該模型實現了課堂教學評估從「事後滯後」向「動態共生」的範式轉變，同時有效提升教師的「實踐知識」與「教育同理心」。本研究旨在建立以人為核心的評估體系，推動教育在科學與教育融合背景下實現培育全面發展個體這一根本使命。

**關鍵詞：**具身認知；多模態數據採集；課堂教學評價；數智化轉型

---

1. 吳治鋒，廣東技術師範大學在讀碩士。