Artificial Intelligence Education Studies
Volume 1, Issue 3, 2025, Page 83-96
https://doi.org/10.6914/aiese.010307

Artificial Intelligence Education Studies

From Framework to Evidence: Testing Explainable AI Feedback for Leadership Learning in Collaborative VR under the C^{2}L-AI Model

Yahan LI
Shandong Vocational University of Foreign Affairs

Abstract
This study advances a testable account of human-AI collaborative competence via the C²L-AI framework. We integrate Activity Theory, Distributed Cognition, and sociomateriality to reconceptualize collaboration, communication, and leadership alongside AI Interaction Competence. We operationalize constructs with an Evidence-Centered Design (ECD) multimodal matrix spanning NLP, social/epistemic network analysis, and VR behavioral analytics. To generate causal evidence, we propose a three-arm randomized controlled trial in a multi-user VR leadership task comparing Explainable AI (XAI) feedback, standard feedback, and no feedback. We hypothesize that XAI yields greater gains in leadership, communication, and AI interaction competence, mediated by improvements in team cognitive architecture (shared mental models, transactive memory). The work offers a unified theory, measurable indicators, and an empirical pathway for designing effective, ethical human-AI learning systems.

Keywords Human-AI collaboration; Explainable AI; Virtual Reality; Leadership learning; Evidence-Centered Design

1 Introduction

1.1 The AI Paradox and the Rise of Human-Centric Skills

The contemporary educational and professional landscape is undergoing a profound transformation driven by rapid advancements in Artificial Intelligence (AI). As AI systems demonstrate increasing proficiency in executing complex analytical processes and automating routine tasks, a seemingly counterintuitive phenomenon is emerging: a critical surge in the demand for uniquely human skills. This ' "AI paradox" ' highlights that as machines become more capable, the distinct abilities that differentiate human intelligence—such as nuanced social interaction, emotional depth, and contextual understanding—gain unprecedented value. These competencies, prominently including collaboration, communication, and leadership, are inherently relational, adaptable to dynamic contexts, and deeply rooted in the human experience.

The growing emphasis on these human-centric skills is not merely a preference but an escalating economic and societal imperative. Research from institutions such as the McKinsey Global Institute indicates that while the demand for advanced IT and data analytics skills is rising, there is an equally pressing shortage of professionals equipped with critical thinking, creativity, and the ability to teach and train others—skills that are difficult for current AI to authentically replicate. This global skill shift compels educational systems to re-evaluate their priorities and methodologies, moving beyond training individuals merely to *use* AI towards cultivating their capacity to work *synergistically* with it, leveraging its computational power while amplifying their own distinctly human contributions.

© The Authors 2025. Article History Received: June 8, 2025; Accpted: July 9, 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. This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License, which permits non-commercial use, sharing, adaptation, distribution, and reproduction in any medium or format, provided the original author(s) and source are credited, and derivative works are licensed under the same terms.

1.2 Beyond "Automation Complementarity" to "Human-AI Co-evolution"

The discourse surrounding AI's impact often centers on a model of ' "automation complementarity," ' where AI automates certain tasks, thereby freeing humans to focus on more complex, creative, and interpersonal endeavors. In this view, human skills and AI capabilities have a static, synergistic relationship. However, this model is insufficient to capture the dynamic and transformative nature of the human-AI partnership. A more forward-looking perspective, and the one adopted in this paper, is that of ' "human-AI co-evolution." '

This co-evolutionary model posits a dynamic, reciprocal feedback loop. As humans interact with increasingly sophisticated AI, their own skills are not just applied but are actively reshaped and refined. For instance, collaboration strategies must adapt to incorporate AI agents as team members, and communication practices evolve to include effective prompting and critical interpretation of AI-generated content. Concurrently, as developers and educators gain a deeper understanding of these new human practices, they can design AI tools that better support and augment these evolving capabilities. This reframes the educational challenge from a reactive, one-time upskilling event to the cultivation of a continuous, adaptive capacity to learn and grow *with* AI. This vision elevates the research agenda from mere adaptation to a proactive shaping of our collective future with intelligent technologies.

1.3 The Research Gap: The Need for an Integrated and Testable Framework

Despite widespread recognition of the importance of human skills in the AI era, a critical gap persists in the academic literature: the lack of an integrated, theoretically-grounded, and empirically-testable framework that defines and operationalizes the competencies required for effective human-AI collaboration. Existing research often treats ' "human skills" ' (e.g., collaboration) and ' "AI tools" ' as separate and distinct entities, failing to adequately theorize their deep and constitutive entanglement in authentic learning and work contexts. Consequently, educators and technology designers lack a coherent model to guide instructional design, assessment practices, and the development of next-generation educational technologies. Without such a framework, efforts to cultivate these crucial skills risk being ad-hoc, difficult to measure, and disconnected from the sociotechnical reality of modern work.

1.4 The Current Study and Its Contributions

This paper aims to fill this critical research gap by proposing and providing an initial validation pathway for a novel theoretical model. It transforms a narrative exploration of the topic into a rigorous, evidence-based academic argument structured around three core contributions:

C1: A Novel Theoretical Framework. First, we propose the Collaborative Competence in Human-AI Learning (C^2L-AI) framework, a multi-dimensional sociotechnical model that integrates core human skills, AI Interaction Competence, and team cognitive architecture to provide a comprehensive new lens for understanding human-AI synergy.

C2: A Rigorous Assessment Methodology. Second, we operationalize this framework by presenting a *multi-modal assessment matrix grounded in Evidence-Centered Design (ECD)*. This matrix details how the framework's abstract constructs can be systematically measured using advanced learning analytics techniques, providing a blueprint for future empirical research.

C3: Causal Empirical Evidence. Third, we provide an initial validation of the framework's utility by designing and proposing a *randomized controlled trial that uses Virtual Reality (VR) and Explainable AI (XAI)* to test the causal effect of interpretable AI feedback on the acquisition of leadership skills in a collaborative context.

By integrating theory, methodology, and a plan for empirical validation, this paper seeks to move the field from broad assertions about the importance of human skills to a scientifically grounded understanding of how to define, measure, and cultivate them in the increasingly complex world of human-AI co-evolution.

2 Theoretical Framework: The C^2L-AI Model

To build a framework capable of capturing the complexity of human-AI collaborative learning, it is necessary to move beyond traditional learning models that locate cognition solely within the individual mind. The introduction of a sophisticated AI agent into a learning group fundamentally alters the unit of analysis, demanding an integrated theoretical perspective that can account for social, cultural, technical, and cognitive factors simultaneously. This section synthesizes multiple theoretical lenses to construct a unified sociotechnical foundation upon which the C^2L-AI framework is built.

2.1 Toward a Unified Sociotechnical Theory of Human-AI Learning

The proposed theoretical foundation is not a simple aggregation of existing theories but a nested, multi-level synthesis where each theory addresses a different analytical challenge. Activity Theory provides the overarching *structure* of the system, Distributed Cognition explains the *cognitive processes* within that structure, and Sociomateriality offers the *ontological stance* that prevents an artificial separation of human and technical elements. This integrated approach provides a powerful toolkit for analyzing human-AI collaboration with unprecedented depth.

2.1.1 Activity Theory (AT): A Macro-Analytic Framework

Activity Theory (AT), originating from the sociocultural psychology of Vygotsky and later expanded by Engeström, provides a macro-level framework for analyzing the learning environment as a complete *activity system*. AT posits that human action is goal-driven, tool-mediated, and unfolds within a collective, sociocultural context. Engeström's second-generation model identifies six core components: the *Subject* (the learner or group), the *Object* (the learning goal or motive), *Mediating Artifacts* (tools used to act on the object), *Rules* (norms governing the activity), the *Community* (the social group sharing the object), and the *Division of Labor* (the distribution of tasks and power).

Within a human-AI learning context, this model offers a powerful analytic lens. The learning team is the Subject, and their learning goal is the Object. Crucially, AI is conceptualized not as a passive tool but as a potent *Mediating Artifact* that actively reshapes the entire system. The introduction of an AI agent can alter the Rules of engagement (e.g., how information is sought), the interaction patterns of the Community, and, most importantly, the Division of Labor between human learners and the AI itself. AT thus allows for an analysis of the system as a whole, focusing on the tensions and contradictions that arise and drive its development.

2.1.2 Distributed Cognition (DC): A Micro-Analytic Framework

While AT describes the ‘ “what” ‘ and ‘ “why” ‘ of the activity system, Distributed Cognition (DC) provides a micro-level perspective on ‘ “how” ‘ cognitive processes unfold within it. Pioneered by Edwin Hutchins, DC theory argues that cognition is not confined to the individual's mind but is *distributed* across the entire system, which includes team members, artifacts (like AI), and the shared environment. In this view, a group of people working together with their tools constitutes a single, integrated cognitive system.

This perspective is essential for understanding how a human-AI team performs cognitive tasks such as memory, problem-solving, and decision-making as a unified entity. The AI is not merely an external resource but an integral part of the team's cognitive architecture. For example, it can function as the team's external memory, a computational engine for complex analysis, or a reasoning partner for simulating outcomes. The team's overall cognitive performance is therefore an emergent property of the dynamic coupling and coordination between human cognitive capabilities and the AI's computational functions.

2.1.3 Sociomateriality: An Entangled Ontology

Finally, the theory of sociomateriality provides an ontological foundation that avoids a simplistic dualism between ‘ “human” ‘ and ‘ “AI.” ‘ Advanced by scholars like Wanda Orlikowski, this perspective holds that the social (human interaction, meaning-making) and the material (AI algorithms, interfaces, data structures) are not separate entities that interact but are ‘ “constitutively entangled.” ‘ Learning practices emerge from this inseparable entanglement, meaning that one cannot analyze ‘ “human skills” ‘ in isolation from the ‘ “AI tools” ‘ through which they are enacted. The AI is not just a tool *used by* humans; it actively participates in and shapes the very practices of collaboration, communication, and leadership. For example, the design of an AI's dashboard (material) shapes the leader's situational awareness (social), and the leader's queries (social) in turn shape the data presented by the AI (material).

2.2 Dimension 1: Core Human Skills (Re-contextualized)

The first dimension of the C^{2}L-AI framework redefines core human skills, moving them from abstract, individual traits to situated practices enacted within the sociotechnical system described above.

Collaboration: Defined as the process of *co-regulating* the human-AI activity system to achieve a shared *Object* (learning goal). This extends beyond simple teamwork to include negotiating the *Division of Labor* between human and AI agents and resolving contradictions that emerge within the system.

Communication: Defined as the set of practices for establishing and maintaining shared understanding, or *inter-subjectivity*, within a distributed cognitive system. This encompasses not only human-human dialogue but also the effective exchange of information with an AI to build a common ground of knowledge.

Leadership: Defined as the function of *ethically steering* the activity system through complexity and uncertainty. This involves managing the responsible integration of AI tools, ensuring their use aligns with ethical norms, and fostering a team culture of critical engagement with and adaptation to AI-generated insights.

2.3 Dimension 2: AI Interaction Competence (AIC)

This dimension formalizes the crucial meta-skill of engaging effectively with AI systems, drawing upon and extending research in AI literacy. AIC is defined as the capacity to effectively and critically manage the human-AI interface within the activity system. It comprises three interrelated sub-constructs:

Operational Proficiency: The technical ability to use AI tools to achieve specific goals. This includes skills such as effective prompt engineering, understanding the functionality of different AI models, and correctly interpreting their output formats.

Critical Evaluation: The ability to assess AI-generated outputs for accuracy, relevance, and potential bias, while understanding the AI's operational limitations and inherent uncertainties. This requires learners to treat AI outputs not as infallible truths but as information sources that require verification and critical integration.

Interactional Attunement: A novel socio-cognitive construct representing the ability to adapt one's innate human ‘ “interactional instincts” ‘ to the non-human nature of an AI agent. This involves avoiding inappropriate anthropomorphism, building an accurate mental model of the AI's capabilities and reasoning processes, and developing new forms of ‘ “dialogue” ‘ suited to a non-human intelligence.

2.4 Dimension 3: Team Cognitive Architecture

This dimension, drawn directly from DC and team science literature, describes the emergent cognitive properties of the human-AI team as a whole. These are not properties of any single individual but reside in the patterns of interaction among all system components, including the AI.

Shared Mental Models (SMMs): The degree of overlap in understanding among human team members regarding the task, the team's capabilities, the environment, and, critically, the AI's role, strengths, and limitations. A meta-analysis by DeChurch and Mesmer-Magnus (2010) established a strong link between SMMs and both team processes and performance, underscoring their importance as a team-level construct. In a human-AI team, a high-quality SMM means all members have a consistent and accurate understanding of what tasks are best suited for the AI versus human members.

Transactive Memory Systems (TMS): The shared, distributed knowledge of ‘ “who knows what” ‘ within the team. As originally theorized by Wegner, a TMS allows a group to collectively store and retrieve more information than any individual could alone. The C^2L-AI framework extends this concept to include an accurate, shared understanding of the AI's specific knowledge domains, data access, and processing capabilities. An effective human-AI TMS enables the team to efficiently direct queries to the correct expert, whether that expert is a human or the AI.

2.5 Counterfactual Distinction and Novelty of the C^2L-AI Framework

The C^2L-AI framework's primary contribution lies in its integrated, sociotechnical approach, which distinguishes it from prior models in related fields:

Distinction from Computer-Supported Collaborative Learning (CSCL): While CSCL research has a long tradition of studying technology-mediated collaboration, it has often treated technology as a communication channel or a contextual factor. C^2L-AI, informed by AT and DC, explicitly models the AI as an active, cognitive agent within the learning system, fundamentally altering the system's dynamics rather than just mediating them.

Distinction from Team Science: The framework is deeply indebted to concepts like SMMs and TMS from team science. However, these models were developed for and tested on human-only teams. C^2L-AI's novelty is in its explicit extension of these constructs to hybrid human-AI teams, which introduces unique challenges such as the inherent opacity of the AI's ‘ “mental model” ‘ and the need for new interaction protocols.

Distinction from AI Literacy: Typical AI literacy frameworks tend to focus on the knowledge and skills of an individual user. C^2L-AI situates the crucial individual skill of AI Interaction Competence (AIC) within a broader collaborative system, linking it directly to team-level processes (SMMs, TMS) and performance outcomes (effective collaboration, communication, and leadership). It moves the focus from individual knowledge *about* AI to collaborative performance *with* AI.

Table 1: The C²L-AI Framework: Constructs, Definitions, and Theoretical Links

Dimension	Construct	Definition	Theoretical Links
1. Core Human Skills (Re-contextualized)	Collaboration	The ability to co-regulate interaction processes, negotiate roles, and resolve contradictions within a human-AI activity system to achieve a shared objective.	Activity Theory, CSCL, Teamwork Models (e.g., Salas et al.)
	Communication	The set of practices enacted to establish and maintain intersubjectivity (shared understanding) within a distributed cognitive system, encompassing both human-human and human-AI information exchange.	Distributed Cognition, CSCL, Sociomateriality
	Leadership	The function of ethically steering the activity system, managing the integration of AI tools, and fostering a team culture of critical engagement and adaptation.	Activity Theory, Teamwork Models
2. AI Interaction Competence (AIC)	Operational Proficiency	The technical ability to effectively use AI tools to achieve sub-goals (e.g., prompt engineering, interpreting basic outputs).	AI Literacy
	Critical Evaluation	The ability to assess AI outputs for accuracy, relevance, and potential bias, and to understand the AI's operational boundaries and limitations.	Critical Thinking Frameworks, Explainable AI
	Interactional Attunement	The socio-cognitive ability to adapt innate ' "interactional instincts" ' to the non-human characteristics of AI, avoiding improper anthropomorphism and building an accurate mental model of the agent.	Human-Computer Interaction, AI Literacy
3. Team Cognitive Architecture	Shared Mental Models	The degree of overlapping understanding among human members, and as represented in the AI's knowledge base, regarding the task, team, and tools.	Team Cognition, Distributed Cognition
	Transactive Memory Systems	The shared, distributed knowledge of ' "who knows what," ' including an accurate understanding of the AI's specific expertise and data access capabilities.	Team Cognition, PISA Collaborative Problem Solving

3 Assessment Methodology: A Multi-modal, Evidence-Centered Approach

To measure the complex, multi-dimensional constructs defined in the C²L-AI framework, traditional assessment methods that rely on self-reports or final products are insufficient. Such methods fail to capture the dynamic, process-oriented nature of skills like collaboration and leadership as they unfold in real-time interactions. This section proposes an innovative, multi-modal assessment methodology grounded in learning analytics and structured by the rigorous principles of Evidence-Centered Design (ECD). The goal is to mine the rich digital traces left by learners in technology-enhanced environments to generate valid, process-based evidence of their competencies. This approach represents a fundamental shift from *assessment of learning* to *assessment for learning*, where the data stream can provide continuous, formative feedback that is integrated into the learning journey itself.

3.1 Grounding Assessment in Evidence-Centered Design (ECD)

ECD is a systematic framework for designing assessments that explicitly links claims about student proficiency to observable evidence, ensuring validity from the outset of the design process. It provides the coherent argumentative structure for our assessment plan by connecting three core models:

Competency Model: This model defines the knowledge, skills, and abilities (KSAs) to be assessed. In this study, the C²L-AI framework itself (Section 2) serves as the Competency Model, providing a detailed specification of the nine target constructs.

Evidence Model: This model specifies which observable learner behaviors or work products (i.e., digital traces) can serve as evidence for the competencies defined in the Competency Model. It articulates the logic for why a particular action (e.g., a student questioning an AI's output) is indicative of a particular skill (e.g., Critical Evaluation).

Task Model: This model describes the design of tasks and scenarios (e.g., a collaborative VR crisis simulation) that are specifically engineered to elicit the behaviors identified in the Evidence Model, providing learners with the opportunity to demonstrate their competence.

This principled approach ensures that the assessment is not an ad-hoc collection of metrics but a defensible system of inference, where every piece of data is purposefully collected and interpreted as part of a larger validity argument.

3.2 A Multi-modal Assessment Matrix for the C^{2}L-AI Framework

The proposed assessment methodology integrates several cutting-edge learning analytics techniques to triangulate evidence from different data sources, enabling a comprehensive measurement of the C^{2}L-AI constructs. The synergy between these methods allows for a holistic understanding that a single technique could not provide; for example, SNA can identify the *structure* of interaction, NLP can analyze the *quality* of that interaction, and ENA can reveal the *cognitive impact* of the interaction on the team's shared understanding.

3.2.1 Assessing Communication (via Natural Language Processing - NLP)

By analyzing transcripts of text-based chat and spoken dialogue, NLP can provide objective, scalable measures of communication quality.

Clarity and Coherence: NLP models can compute readability scores (e.g., Flesch-Kincaid), measure lexical diversity, and apply text coherence models to quantify the clarity and logical flow of communication.

Argumentation Quality: Advanced argumentation mining techniques can automatically identify argumentative components such as *Claims*, *Warrants*, and *Evidence* within student discussions. Analyzing the frequency, structure, and quality of these components provides a deep measure of critical thinking and persuasive communication.

Social and Emotional Tone: Sentiment analysis can be applied to gauge the emotional polarity (positive, negative, neutral) of messages, providing a proxy measure for empathy and the ability to maintain a positive team climate.

3.2.2 Assessing Collaboration (via Social and Epistemic Network Analysis)

By modeling the patterns of interaction within the team, network analysis can reveal the underlying structure and dynamics of collaboration.

Social Network Analysis (SNA): Using interaction data (e.g., who replies to whom in a chat), SNA can construct a social network of the team. Key metrics such as *network density* (a measure of group cohesion) and *centrality* measures (identifying influential members or peripheral participants) can quantify the structure of participation and information flow.

Epistemic Network Analysis (ENA): ENA moves beyond social ties to model the semantic connections in a team's discourse. By quantifying and visualizing how different concepts are linked in the conversation over time, ENA can reveal the structure of the team's knowledge-building process, showing whether they are engaging in deep, connected reasoning or superficial information exchange.

3.2.3 Assessing Leadership and AIC (via VR Behavioral Analytics)

Multi-user VR environments provide a ' "laboratory in the wild," ' offering high ecological validity while capturing fine-grained behavioral data that is difficult to obtain in other settings.

Decision-Making Under Pressure: The VR simulation can track the choices leaders make in high-stakes, time-sensitive situations, the information they seek before deciding, and the speed and quality of those decisions.

Information Management and AI Interaction: The system can log how frequently a leader consults the AI agent, the nature of their queries (mapping to Operational Proficiency), and whether they challenge or seek verification for AI-provided information (mapping to Critical Evaluation).

Situational Awareness: Advanced techniques such as eye-tracking within the VR headset can provide objective measures of a leader's situational awareness, tracking what information they attend to in the complex virtual environment.

4 The VR-XAI Experiment: A Validation Study

While proposing a theoretical framework and an assessment methodology are crucial first steps, rigorous empirical validation is necessary to establish their scientific merit and practical utility. This section moves from proposal to a detailed experimental design aimed at validating a key causal relationship implied by the C^{2}L-AI framework: the link between the quality of AI interaction and the development of core human skills. Specifically, this study investigates how different types of AI feedback within a collaborative VR leadership training environment affect learning outcomes.

Table 2: Multi-modal Assessment Matrix for the C²L-AI Framework

C²L-AI Construct	Data Source	Primary Analysis Technique	Key Metrics/Indicators	References
Collaboration	Chat logs, shared document edit histories, project management tool data	Social Network Analysis (SNA)	Network density, reciprocity (group cohesion); degree and betweenness centrality (participation & influence).	40
Communication	Transcripts of oral/written dialogue	Natural Language Processing (NLP)	Coherence/consistency scores; identification of argumentative components (claim, evidence); sentiment analysis.	28
Leadership	VR simulation interaction logs (decisions, actions, communication with virtual agents)	Behavioral Analytics	Decision speed & quality under pressure; frequency of strategic communication; gaze patterns (situational awareness).	1
AIC: Operational Proficiency	Command logs, interaction history with AI tools	Frequency/Pattern Analysis	Query efficiency; rate of advanced feature use; error rate in tool usage.	1
AIC: Critical Evaluation	Dialogue transcripts, written reflections	NLP Content Analysis	Frequency of challenging AI outputs; use of verification strategies; explicit mentions of AI limitations/bias.	1
AIC: Interactional Attunement	Dialogue transcripts (human-AI prompts & responses)	Qualitative Coding & NLP	Avoidance of anthropomorphic language; prompt structures built for a machine, not a human; metacognitive statements about the AI's nature.	1
Shared Mental Models	Pre/post-test concept maps; dialogue transcripts	ENA / Semantic Similarity Analysis	Convergence of concept maps over time; high semantic similarity of team terminology (using NLP); strong co-occurrence of key concepts in ENA graphs.	16
Transactive Memory Systems	Communication logs (' "Who asks whom what question?" ')	Social Network Analysis (SNA)	Differentiated network roles; efficient information seeking (queries directed to the correct expert, human or AI).	20

4.1 Research Questions and Hypotheses

The study is designed to answer the following research questions and test the corresponding hypotheses:

RQ1: In a collaborative VR training environment, does AI-driven feedback lead to greater learning gains in leadership competence compared to an absence of feedback?

RQ2: Does Explainable AI (XAI) feedback, which provides causal explanations for performance, lead to significantly greater learning gains and deeper understanding than standard, ' "black-box" ' performance feedback?

H1: Participants in both feedback conditions (Standard and XAI) will demonstrate significantly greater pre-to-post-test learning gains in leadership competence than participants in the no-feedback control group.

H2: Participants in the XAI feedback condition will demonstrate significantly greater pre-to-post-test learning gains in leadership competence, skill transfer, and self-efficacy than participants in the standard feedback group.

4.2 Method and Design

The study will employ a *between-subjects, pre-test/post-test randomized controlled trial design.*

Participants: Participants will be recruited from undergraduate or graduate programs in business, management, or related fields. A priori power analysis will be conducted to determine the necessary sample size to detect a medium effect size with a power of 0.80 and an alpha of 0.05. Participants will be randomly assigned to one of three experimental conditions.

Learning Environment: The experiment will be conducted in a multi-user, collaborative VR leadership simulation. Platforms like Mursion or Talespin provide precedents for such environments, which offer high ecological validity by immersing participants in realistic, dynamic scenarios. The task will be designed as an ill-structured problem (e.g., managing a corporate crisis with incomplete information) that requires the team to collaborate, communicate effectively, and demonstrate leadership to succeed.

Independent Variable (The Manipulation): The sole independent variable is the type of AI feedback provided to the teams at critical decision points during the simulation. The three levels are:

1. Control Group (No Feedback): Teams participate in the VR simulation but receive no AI-generated performance feedback.

2. Standard Feedback Group: After each key decision, teams receive a quantitative, ' "black-box" ' performance score (e.g., ' "Your team's decision-making effectiveness was rated 85 out of 100" ').

3. XAI Feedback Group: Teams receive the same quantitative score accompanied by a concise, causal explanation generated by the AI (e.g., ' "Your team's effectiveness was rated 85/100 because you rapidly identified the key information, but you failed to adequately consider the long-term risks of Option B" '). This manipulation directly tests the impact of AI explainability on learning and skill acquisition.

4.3 Procedure

The experimental procedure will be standardized across all conditions, with the exception of the feedback manipulation:

1. Informed Consent and Pre-Test: After providing informed consent, all participants will complete a battery of pre-test measures to establish a baseline for their leadership, communication, and AIC skills, using a subset of the instruments from the multi-modal assessment matrix (Table 2).

2. Randomization: Participants will be randomly assigned to teams, and teams will be randomly assigned to one of the three experimental conditions.

3. VR Intervention: Each team will engage in the same collaborative VR leadership simulation. The AI will play the role of a data analyst or advisor within the simulation. At predefined intervals, the two experimental groups will receive their respective types of AI feedback.

4. Post-Test: Immediately following the intervention, all participants will complete a post-test assessment battery that is parallel in structure and difficulty to the pre-test, measuring learning gains. Additional measures will assess skill transfer to a novel but related problem, as well as self-efficacy and cognitive load.

4.4 Measures and Analysis Plan

Dependent Variables: The primary dependent variable is the *learning gain in leadership competence*, calculated as the difference between post-test and pre-test scores derived from the VR behavioral analytics. Secondary dependent variables include learning gains in communication and AIC, and self-reported measures of self-efficacy and cognitive load.

Measurement Invariance: A critical preliminary step is to establish the validity of the assessment tools across the experimental groups. Before testing the main hypotheses, *Multi-Group Confirmatory Factor Analysis (MG-CFA)* will be conducted on the measurement models for the key constructs. This statistical procedure tests whether the assessment instruments are measuring the same underlying psychological constructs in the same way across the control, standard, and XAI groups. Establishing measurement invariance (at the configural, metric, and scalar levels) is a prerequisite for making meaningful and valid comparisons of group means. This step is the methodological linchpin that ensures any observed differences are attributable to the intervention, not to a measurement artifact.

Hypothesis Testing: The primary statistical analysis to test H1 and H2 will be a one-way *Analysis of Covariance (ANCOVA)*. For each key dependent variable (e.g., post-test leadership score), the ANCOVA will compare the means of the three groups while statistically controlling for participants' corresponding pre-test scores, which will be entered as a covariate. This method provides a more precise estimate of the intervention's net effect by accounting for initial differences between participants.

Process Analysis: To explore the mechanisms behind any observed effects, the rich process data from the VR interaction logs and communication transcripts will be analyzed. Techniques such as *sequential analysis* or *Epistemic Network Analysis (ENA)* will be used to compare the collaborative patterns and knowledge-building structures of the teams across the three conditions, investigating *how* XAI feedback may have altered the collaborative process itself.

5 Anticipated Results and Contributions

As the experiment outlined is a proposed study, this section describes the anticipated pattern of results and discusses how these findings would contribute significantly to the field.

5.1 Anticipated Main Effects

The primary hypothesis (H2) predicts a significant main effect of the feedback condition on leadership learning gains. It is anticipated that the ANCOVA will yield a statistically significant difference between the three groups' adjusted post-test means. Subsequent post-hoc comparisons (e.g., using Tukey's HSD) are expected to reveal a specific ordinal pattern: the XAI feedback group's mean score will be significantly higher than the standard feedback group's mean, which in turn will be significantly higher than the control group's mean (XAI > Standard > Control). This outcome would provide strong, causal evidence supporting the core hypotheses, demonstrating not only that AI feedback is beneficial for learning complex skills, but that the *explainability* of that feedback is a critical determinant of its effectiveness.

5.2 Anticipated Process-Level Findings

The analysis of process data is expected to illuminate the mechanisms underlying the main effects. For instance, an ENA of the teams' dialogues is anticipated to show that conversations in the XAI group feature a denser network

Table 3: VR-XAI Experiment Design Summary

Component	Description
Research Questions	What is the causal effect of Explainable AI (XAI) feedback, compared to standard AI feedback and no feedback, on the development of leadership competence in a collaborative VR training environment?
Design	Between-subjects, pre-test/post-test randomized controlled trial.
Independent Variable	Type of AI Feedback (3 levels): 1. Control (No AI feedback), 2. Standard Feedback (Performance score only), 3. XAI Feedback (Performance score + causal explanation).
Dependent Variables	Primary: *Learning gain in Leadership Competence* (measured via VR behavioral analytics). Secondary: *Learning gains in Communication and AIC*; *Self-Efficacy* (questionnaire); *Cognitive Load* (questionnaire).
Task Environment	A multi-user, collaborative VR simulation requiring team-based decision-making under uncertainty and time pressure.
Procedure	1. Pre-test assessment. 2. Random assignment to conditions. 3. VR training intervention. 4. Post-test assessment.
Analysis Plan	1. *Psychometrics*: Use Multi-Group Confirmatory Factor Analysis (MG-CFA) to test for measurement invariance of assessment tools across groups. 2. *Main Analysis*: Use Analysis of Covariance (ANCOVA) to compare post-test scores on dependent variables, with pre-test scores as covariates. 3. *Process Analysis*: Use Sequential Analysis or ENA on interaction logs to explore mechanisms.

of connections between identified problems, evidence considered, and proposed solutions. This would suggest that the causal explanations provided by the AI prompted deeper, more structured reasoning and helped the team build a more robust shared mental model of the task. In contrast, the standard feedback group's network might show a focus on performance scores without a corresponding increase in conceptual linkage, while the control group's network would be the least structured. Such findings would suggest that the superior learning gains in the XAI group are mediated by an improvement in the quality of the team's collaborative knowledge-building process.

5.3 Implications of Anticipated Findings

If the results align with these expectations, they would constitute a major contribution to both theory and practice. They would provide some of the first rigorous, causal evidence that the *manner* in which an AI system communicates its reasoning directly and significantly impacts human skill acquisition in a complex, collaborative domain. This would validate a central tenet of the C^{2}L-AI framework: that the quality of the human-AI interaction, as captured in the AIC dimension (specifically, the AI's capacity to provide explanations that support users' Critical Evaluation), is a critical lever for developing the Core Human Skills dimension (Leadership). The findings would lend strong empirical support to the framework's structure and its utility as a model for designing effective human-AI learning systems.

6 Discussion

The anticipated findings from the VR-XAI experiment hold profound implications for theory, instructional practice, and the design of educational technology. This section interprets these potential results, discusses their broader significance, and acknowledges the study's limitations while charting a course for future research.

6.1 Theoretical Implications

The expected results would provide strong empirical grounding for the C^{2}L-AI framework, particularly by demonstrating the critical, causal link between the AI Interaction Competence (AIC) dimension and the Core Human Skills dimension. The superior performance of the XAI group would suggest that when an AI's feedback is transparent and interpretable, it does more than simply provide information; it scaffolds the user's ability to critically evaluate situations, a key component of both AIC and effective leadership. This supports the framework's core assertion that these competencies are not developed in isolation but are deeply interconnected within a single sociotechnical system.

Furthermore, these findings would lend credence to the concept of ' "human-AI co-evolution." ' A well-designed XAI that provides causal feedback acts as a catalyst in this evolutionary spiral. It enhances human understanding, which leads to better performance and more sophisticated skill application. This, in turn, provides richer data that can be used to further refine the AI's support. This moves beyond a simple view of AI as a tool and positions it as a partner in a developmental process, with explainability being a key mechanism that facilitates this partnership.

6.2 Practical and Design Implications

The practical implications for instructional designers and educational technology developers are direct and actionable. The central message would be that for AI to be a truly effective educational partner, particularly for complex skill development, it is not enough for it to be *correct*; it must be *interpretable*.

For Instructional Design: Educators designing learning experiences that incorporate AI should prioritize activities that require learners to engage with, question, and even challenge AI-generated outputs. The focus should shift from simply using AI to get answers to using AI to refine one's own reasoning and decision-making processes.

For Technology Development: The findings would compel developers of educational AI to move beyond ' "black-box" ' models that deliver only performance scores or recommendations. Instead, they should invest in robust XAI features that provide users with clear, concise, and causal explanations for the AI's judgments. For example, an automated essay-scoring tool should not just provide a grade but should highlight specific rhetorical moves that contributed positively or negatively to the score, linking its evaluation back to the rubric in an understandable way.

6.3 Limitations and Future Research

While the proposed experimental design is rigorous, it is essential to acknowledge its limitations. As a laboratory-based study with a specific population (e.g., university students) and a single task context (a specific VR crisis scenario), the generalizability of the findings to other populations, domains, and real-world settings must be approached with caution. The short-term nature of the intervention also does not allow for conclusions about the long-term retention and transfer of the skills learned.

These limitations point directly to a rich agenda for future research. The experiment should be replicated in diverse settings, such as K-12 education or corporate training, and with different complex tasks like collaborative scientific inquiry or medical diagnosis, to test the robustness of the findings. Future research should also employ longitudinal designs to track the development of C^2L-AI competencies over extended periods, examining how skills are retained and transferred to authentic workplace or academic environments. Furthermore, as the proposed study uses one form of textual XAI, future work should investigate the differential impacts of various XAI modalities (e.g., visual explanations, interactive ' "what-if" ' scenarios) and levels of complexity on learning. Finally, given that this study tests one critical link within the C^2L-AI framework, a broader research program is needed to empirically validate the other proposed relationships, such as how SMMs and TMS mediate the relationship between collaborative processes and team outcomes in human-AI teams.

7 Ethics and Governance

The integration of AI into the development and assessment of human skills, while promising, is laden with ethical challenges. A reactive, post-hoc approach to ethics is insufficient. Instead, responsible innovation demands that ethical principles be proactively embedded into the design, development, and deployment of these systems from their inception. This section outlines a comprehensive governance framework to ensure that human-AI learning systems are fair, transparent, secure, and respectful of human agency.

7.1 A Proactive Framework for Responsible Innovation

The ethical implementation of AI in education requires moving beyond compliance checklists to a culture of deep ethical reflection. The framework proposed here is organized around four core principles: Fairness, Transparency, Privacy, and Human Agency. These principles must be operationalized through concrete processes and technical safeguards throughout the AI system's lifecycle.

7.2 Operationalizing Ethical Principles

The following subsections detail specific, actionable strategies for implementing each ethical principle within the context of systems like the one proposed in this paper.

7.2.1 Fairness and Bias Mitigation

AI models trained on historical data can inadvertently learn, perpetuate, and even amplify existing societal biases related to race, gender, socioeconomic status, and other characteristics. An AI system designed to assess leadership, if trained on biased data, could systematically disadvantage individuals from underrepresented groups, thereby exacerbating educational and professional inequities.

To combat this, a mandatory *AI Bias Audit* must be a standard procedure before any educational AI system is deployed. This audit should be multi-faceted, encompassing a data audit, a model audit, and a pedagogical audit. The data audit involves scrutinizing the training datasets for demographic representativeness and identifying potential sources of historical bias. The model audit consists of conducting performance testing of the AI model across different demographic subgroups to detect any disparate impact, using fairness metrics (e.g., equalized odds) to quantify and report on model fairness. Finally, the pedagogical audit evaluates the underlying instructional logic of the AI; for instance, an adaptive system must be checked to ensure it does not systematically channel certain student groups towards less challenging or less valuable learning pathways.

7.2.2 Transparency and Explainability (XAI)

The ‘ “black box” ‘ nature of many sophisticated AI models is ethically untenable in high-stakes educational contexts. If a student receives a poor assessment from an AI but cannot understand the basis for that judgment, the assessment is pedagogically useless and ethically irresponsible. Therefore, *Explainable AI (XAI) is not merely a desirable feature but an ethical imperative.* Assessment and feedback systems must be designed to provide clear, human-understandable justifications for their outputs. This transparency is fundamental for building trust with students and educators, enabling meaningful reflection, and providing a basis for contesting or appealing an AI's judgment.

7.2.3 Privacy and Security

AI systems for personalized learning and assessment often collect vast quantities of sensitive student data, creating significant privacy risks. To ensure robust data governance, the use of *Data Flow Diagrams (DFDs)* should be standard practice. A DFD is a visualization that maps precisely how student data is collected, processed, stored, and shared. This process helps ensure:

Data Minimization: Only data that is strictly necessary for the stated educational purpose is collected.

Security: Potential vulnerabilities in data storage and transmission are identified and mitigated through measures like encryption and anonymization.

Compliance and Consent: The data flows are transparent to students and parents, facilitating informed consent and ensuring compliance with regulations such as the Family Educational Rights and Privacy Act (FERPA).

7.2.4 Human Agency and Oversight

The goal of AI in education should be to augment human intelligence, not supplant human judgment. This requires a steadfast commitment to the ‘ “human-in-the-loop” ‘ principle, especially for high-stakes decisions. Any AI-generated assessment that could significantly impact a student's academic standing or future opportunities must be subject to review and final approval by a qualified human educator.

Furthermore, institutions must establish a clear, accessible, and fair *Student Appeal Process.* If a student believes an AI's assessment is inaccurate or unjust, they must have a formal recourse to have their case reviewed by human decision-makers. This mechanism is essential for safeguarding student rights and maintaining human accountability at the center of the educational process.

8 Conclusion

The rapid ascent of AI has created a pivotal moment for education, paradoxically illuminating the escalating value of uniquely human skills. This paper has argued that in an era where AI can perform a vast array of analytical tasks, a ‘ “renaissance” ‘ of human competencies—particularly collaboration, communication, and leadership—is a fundamental necessity for individual and societal flourishing. The journey from theorizing a new sociotechnical model of human-AI competence to designing a rigorous methodology for its assessment and a causal experiment for its validation converges on a singular theme: the future of education must be profoundly human-centric, even as it becomes increasingly technologically advanced.

8.1 Summary of Contributions

This paper has sought to advance the field by making four primary contributions. First, through a unified theoretical synthesis integrating Activity Theory, Distributed Cognition, and Sociomateriality, it offers a novel sociotechnical framework for analyzing human-AI learning that moves beyond simplistic dichotomies of human versus machine. Second, it proposes the innovative C^2L-AI framework, a multi-dimensional model that systematically defines the core human skills, AI interaction competencies, and team cognitive architecture essential for success in human-AI collaborative environments. Third, it presents a rigorous multi-modal assessment methodology; grounded in Evidence-Centered Design, this operational and data-driven methodology leverages NLP, SNA, ENA, and VR behavioral analytics to measure the dynamic processes of collaboration, moving assessment from a summative event to a formative, integrated part of learning. Finally, it provides both an empirical pathway and an ethical guide by offering a clear, rigorous experimental design for empirically testing the framework's core tenets and a comprehensive ethical implementation checklist to guide the responsible design and deployment of educational AI systems.

8.2 Human Skill Resilience in an Age of Co-evolution

The ultimate argument of this paper returns to the theme of ‘ “human-AI co-evolution.” ‘ The final goal of this line of inquiry is to foster what can be termed ‘ “Human Skill Resilience” ‘: a deeply ingrained capacity to continuously learn, adapt, collaborate effectively, communicate with clarity and empathy, and lead with ethical integrity, irrespective of how AI technologies transform the world. This resilience, rooted in well-developed and adaptable human skills, is the cornerstone for navigating an uncertain future and ensuring that humanity remains the architect of its own destiny. In this co-evolutionary future, the most critical educational outcome is not merely to adapt *to* AI, but to engage in a dynamic process where human skills become more sophisticated through interaction with intelligent systems, and AI tools, in turn, are designed to better support these evolving human capabilities.

8.3 Future Research Directions

The intersection of AI and human skill development is a nascent and fertile ground for future inquiry. Building on the work presented here, several research directions are particularly pressing. There is a critical need for longitudinal research to understand the long-term impact of AI-augmented learning environments on the development, retention, and transfer of collaboration, communication, and leadership skills. Further research is crucial in developing and validating robust, context-aware, and practical XAI methods specifically designed for the assessment of complex human skills in educational settings. It is also essential to investigate the most effective models for human-AI hybrid feedback and instruction, balancing the scalability of AI with the pedagogical richness of human mentorship. Moreover, exploring the nuances of C^2L-AI competencies in diverse human-AI team configurations and across different cultural contexts will yield valuable insights into the generalizability of these models. A global perspective requires focused research and international collaboration to address the challenges of scaling equitable access to high-quality AI educational tools and implementing robust ethical frameworks in under-resourced regions and diverse educational systems. Finally, understanding how AI can support the continuous adaptation of human skills throughout an individual's career is increasingly important for lifelong learning in a rapidly changing world of work.

To Cite This Article Yahan LI. (2025). From Framework to Evidence: Testing Explainable AI Feedback for Leadership Learning in Collaborative VR under the C^2L-AI Model. *Artificial Intelligence Education Studies*, *1*(3), 83-96. https://doi.org/10.6914/aiese.010307

Reference

[1] Baker, R. S., & Hawn, A. (2021). Algorithmic bias in education. *International Journal of Artificial Intelligence in Education, 32*(4), 1052–1092. https://doi.org/10.1007/s40593-021-00285-1

[2] Bisantz, A. M., & Roth, E. M. (2007). Analysis of cognitive work. *Reviews of Human Factors and Ergonomics, 3*(1), 1–43. https://doi.org/10.1177/1557234X0700300101

[3] Bolstad, C. A., & Endsley, M. R. (1999). Shared displays and team performance. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 43*(3), 205–209. https://doi.org/10.1177/154193129904300301
[4] Cannon-Bowers, J. A., & Salas, E. (2001). Reflections on shared cognition. *Journal of Organizational Behavior, 22*(2), 195–202. https://doi.org/10.1002/job.82
[5] Cannon-Bowers, J. A., Salas, E., & Converse, S. (1993). Shared mental models in expert team decision making. In N. J. Castellan, Jr. (Ed.), *Individual and group decision making: Current issues* (pp. 221–246). Lawrence Erlbaum Associates, Inc.
[6] Cooke, N. J., Gorman, J. C., Duran, J. L., & Taylor, A. R. (2007). Team cognition in experienced command-and-control teams. *Journal of Experimental Psychology: Applied, 13*(3), 146–157. https://doi.org/10.1037/1076-898X.13.3.146
[7] DeChurch, L. A., & Mesmer-Magnus, J. R. (2010a). The cognitive underpinnings of effective teamwork: A meta-analysis. *Journal of Applied Psychology, 95*(1), 32–53. https://doi.org/10.1037/a0017225
[8] DeChurch, L. A., & Mesmer-Magnus, J. R. (2010b). Measuring shared team mental models: A meta-analysis. *Group Dynamics: Theory, Research, and Practice, 14*(1), 1–14. https://doi.org/10.1037/a0017455
[9] Engeström, Y. (1999). Activity theory and individual and social transformation. In Y. Engeström, R. Miettinen, & R.-L. Punamäki (Eds.), *Perspectives on activity theory* (pp. 19–38). Cambridge University Press.
[10] Favero, F., Gessler, D., & Dragoni, M. (2025). *Argument mining in student essays with small large language models.* arXiv preprint arXiv:2502.14389. https://arxiv.org/abs/2502.14389
[11] Hirschfeld, G., & von Brachel, R. (2014). Improving multiple-group confirmatory factor analysis in R A tutorial in measurement invariance with continuous and ordinal indicators. *Practical Assessment, Research, and Evaluation, 19*(7). https://doi.org/10.7275/qazy-2946
[12] Hollan, J., Hutchins, E., & Kirsh, D. (2000). Distributed cognition: Toward a new foundation for human-computer interaction research. *ACM Transactions on Computer-Human Interaction,* 7(2), 174 196. https://doi.org/10.1145/353485.353487
[13] Hollingshead, A. B. (2001). Cognitive interdependence and convergent expectations in transactive memory. *Journal of Personality and Social Psychology, 81*(6), 1080 1089. https://doi.org/10.1037/0022-3514.81.6.1080
[14] Hutchins, E. (1995). *Cognition in the wild.* MIT Press.
[15] Hutchins, E. (2006). The new distributed cognition. *Pragmatics & Cognition, 14*(3), 435–446.
[16] Johri, A. (2011). The socio-materiality of learning practices and implications for the field of learning technology. *Research in Learning Technology, 19*(3). https://doi.org/10.3402/rlt.v19i3.17139
[17] Khosravi, H., Shum, S. B., Chen, G., Conati, C., Tsai, Y. S., Kay, J., Knight, S., Martinez-Maldonado, R., Sadiq, S., & Gašević, D. (2022). Explainable artificial intelligence in education. *Computers and Education: Artificial Intelligence, 3*, 100074. https://doi.org/10.1016/j.caeai.2022.100074
[18] Klimoski, R., & Mohammed, S. (1994). Team mental model: Construct or metaphor? *Journal of Management, 20*(2), 403 437. https://doi.org/10.1177/014920639402000206
[19] Lawrence, J., & Reed, C. (2020). Argument mining: A survey. *Computational Linguistics, 45*(4), 765 818. https://doi.org/10.1162/coli_a_00364
[20] Lewis, K. (2003). Measuring transactive memory systems in the field: Scale development and validation. *Journal of Applied Psychology, 88*(4), 587 604. https://doi.org/10.1037/0021-9010.88.4.587
[21] Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1 16. https://doi.org/10.1145/3313831.3376727
[22] Mcleod, S. (2024, August 8). *Cultural-historical activity theory*. Simply Psychology. Retrieved September 4, 2025, from https://www.simplypsychology.org/cultural-historical-activity-theory.html
[23] Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (2003). On the structure of educational assessments. *Measurement: Interdisciplinary Research and Perspectives, 1*(1), 3–62. https://doi.org/10.1207/S15366359MEA0101_02
[24] Mohammed, S., & Dumville, B. C. (2001). Team mental models in a team knowledge framework: Expanding theory and measurement across disciplinary boundaries. *Journal of Organizational Behavior, 22*(2), 89–106. https://doi.org/10.1002/job.86
[25] Nardi, B. (1996). *Activity theory and human-computer interaction.* In B. Nardi (Ed.), Context and Consciousness: Activity Theory and Human-Computer Interaction. MIT Press. Retrieved September 4, 2025, from https://ics.uci.edu/~corps/phaseii/nardi-ch1.pdf
[26] Orlikowski, W. J., & Scott, S. V. (2008). Sociomateriality: Challenging the separation of technology, work and organization. *The Academy of Management Annals, 2*(1), 433–474. https://doi.org/10.5465/19416520802211644

[27] Pöysä-Tarhonen, J., Awwal, N., Häkkinen, P., & Shaffer, D. W. (2024). Understanding knowledge building processes using epistemic network analysis and sequential analysis. *Proceedings of the 18th International Conference of the Learning Sciences - ICLS 2024*, 399–400. International Society of the Learning Sciences.

[28] Rienties, B., Giesbers, B., Tempelaar, D., & Lygo-Baker, S. (2013). How social network analysis can be used to monitor online collaborative learning and guide an informed intervention. *Educational Technology & Society, 16*(3), 234–247.

[29] Salas, E., Sims, D. E., & Burke, C. S. (2005). Is there a "big five" in teamwork? *Small Group Research, 36*(5), 555 599. https://doi.org/10.1177/1046496405277134

[30] Shaffer, D. W., Collier, W., & Ruis, A. R. (2016). A tutorial on epistemic network analysis: Analyzing the structure of connections in cognitive, social, and interaction data. *Journal of Learning Analytics, 3*(3), 9 45. https://doi.org/10.18608/jla.2016.33.3

[31] Stahl, G. (2002). A theory of group cognition in CSCL. *Proceedings of the Conference on Computer Support for Collaborative Learning: Foundations for a CSCL Community*, 73–82. https://doi.org/10.3115/1658616.1658628

[32] The Decision Lab. (n.d.). *Distributed Cognition.* Retrieved September 4, 2025, from https://thedecisionlab.com/reference-guide/neuroscience/distributed-cognition

[33] Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*(1), 4–70. https://doi.org/10.1177/109442810031002

[34] Wegner, D. M. (1987). Transactive memory: A contemporary analysis of the group mind. In B. Mullen & G. R. Goethals (Eds.), *Theories of group behavior* (pp. 185 208). Springer-Verlag.

[35] Yuan, J., Liu, C., & Zhai, X. (2024). A meta-analysis of the effect of explainable artificial intelligence on adaptive learning. *Journal of Computer-Aided Design & Computer Graphics, 36*(10), 10. https://doi.org/10.3969/j.issn.1009-8097.2024.10.010

Editor *Changkui LI wtocom@gmail.com*