SAGE: A Compositional Multi-Agent LLM Framework with Pedagogical Reasoning for Structured Collaborative Problem Solving

Van-Khanh Tran[1,2*], **Van-Khai Dang**[3†], **Duc-Huy Nguyen**[3†]

[1]Institute of Artificial Intelligence,
Thai Nguyen University of Information and Communication Technology, Thai Nguyen, Vietnam
[2]FPT Technology Research Institute,
Education Zone, Hoa Lac High Tech Park, Km29 Thang Long Boulevard, Hoa Lac Commune, Hanoi, Vietnam
[3]Institute of Artificial Intelligence,
VNU University of Engineering and Technology, Hanoi, Vietnam

Abstract

While AI can simulate virtual classrooms, effective collaborative learning requires both dynamic interaction and a well-structured pedagogical plan. To address this, we introduce SAGE (Scaffolded Agent-Guided Education), a novel, compositional two-phase framework. First, a planning module automatically generates an optimized pedagogical scenario using a dedicated team of agents. Second, this scenario is used to configure a conversation module, where autonomous agents engage a student in a structured, real-time dialogue. This approach ensures that dynamic, multi-agent interactions are grounded in a pedagogically sound foundation. We evaluate SAGE through simulation and a study with real students. Results show improved performance against a next-speaker prediction baseline (achieving a 72.13% win rate) and demonstrate effective group dynamics. Specifically, our study with students reveals high role adherence from AI agents, a balanced progression between task-oriented and socio-emotional interactions, and a clear scaffolding effect where instructional support fades as learner autonomy increases. Our findings highlight the significant potential of synergizing automated instructional design with autonomous conversational execution for collaborative learning.

Introduction

While collaborative learning offers rich opportunities for developing communication and reasoning skills, students often struggle with unstructured group work, particularly in complex domains like mathematical problem-solving. This challenge of effective group *execution* is well-known. However, a parallel and often overlooked challenge lies in the *instructional design* itself: crafting high-quality, structured lesson plans is time-consuming, requires significant pedagogical expertise, and is difficult to personalize at scale. Within this context, we focus on collaborative problem-solving (CPS), a specific subset of collaborative learning where peers apply shared knowledge to solve a structured problem, a process that is highly amenable to scaffolding.

*Correspondence: tvkhanh@ictu.edu.vn

†These authors contributed equally.

The recent evolution of Large Language Models (LLMs) presents a unique opportunity to tackle both of these challenges. To this end, we introduce SAGE (Scaffolded Agent-Guided Education), a multi-agent framework designed to holistically address both pedagogical planning and collaborative execution. Building upon prior work in automated instructional design, SAGE operates through a compositional, two-phase architecture:

1. **Automated Planning:** An offline phase where a dedicated team of AI agents automatically generates and refines a detailed pedagogical scenario. This ensures the foundational lesson plan is pedagogically sound.
2. **Real-time Execution:** An interactive phase where the resulting *Optimized Pedagogical Scenario* is used to configure a team of conversational agents. These agents then engage a human student in a dynamic, collaborative dialogue.

We evaluate SAGE through experiments involving Vietnamese high-school students working on mathematical modeling tasks. Results show that the system improves group coordination, diversifies student-agent interaction, and enhances the depth of problem understanding. Our contributions are fourfold:

- **Compositional two-phase architecture:** A framework that uniquely integrates an automated lesson-planning module with a real-time, autonomous conversational system, bridging the gap between instructional design and execution.
- **Automated multi-agent planning:** An adaptation and integration of a planning process that generates optimized pedagogical scenarios to scaffold the subsequent collaborative dialogue.
- **Proactive turn-taking mechanism:** An implementation of a self-selecting turn-taking process within the conversational phase to create natural and educationally-aligned interactions.
- **Comprehensive empirical validation:** A study including both simulation-based benchmarks and experiments with real students, demonstrating significant gains in pedagogical alignment and learner engagement.

Related Work

LLMs for Education

The release of ChatGPT in 2022 introduced a new era in education, shifting from traditional NLP to powerful transformer-based LLMs. Today, these models are widely accessible, enabling automated content creation, real-time feedback and grading at scale, truly personalized learning experiences (Wang et al. 2024). LLMs can role-play historical figures or conversational partners to foster immersive, engaging lessons (Zhu et al. 2025). Researchers even use LLMs to simulate student behavior, comparing their error rates on multiple-choice questions to those of real learners, to generate high-quality assessments (Liu, Sonkar, and Baraniuk 2025).

One-to-one Tutoring

One-to-one tutoring using AI systems, especially those powered by LLMs, leverages various pedagogical strategies to enhance learning outcomes (Gousopoulos 2024), (Razafinirina, Dimbisoa, and Mahatody 2024). While one-to-one tutoring offers personalized attention, it faces challenges in simulating the full spectrum of classroom interactions. One-to-one settings often miss peer learning opportunities, which are crucial for social development and collaborative skills. In contrast, traditional classrooms foster peer interactions that enhance learning through discussion and shared problem-solving. These limitations highlight the need for a more comprehensive approach to simulate realistic learning experiences.

Virtual Classroom – Collaborative Learning

Multi-agent Systems (MAS). In a virtual classroom context, agents can be designed with various roles such as classmates or teachers, collaborating with real students toward shared learning goals. MAS based on Large Language Models (LLMs) (Wang, Zhao, and Liu 2024) have emerged as a potential solution to this challenge, thanks to their capabilities in reasoning, decision-making, and flexible coordination among agents.

Turn-takings in Multi-Party Conversations. Studies such as SimClass (Zhang et al. 2024) and MathVC (Yue et al. 2025) have proposed Next-Speaker Prediction, an approach to managing turn-taking. This method is based on the history and role descriptions of agents to select the most suitable agent to talk. However, this approach leaves agents in a passive position when they are selected by another manager agent. In reality, when people talk to each other, they think independently before speaking. Therefore, a more comprehensive solution is needed to simulate this multi-participant conversation to increase the naturalness of communication.

Recent work (Liu et al. 2025) introduced the *Inner Thoughts* framework, which enables proactive conversational agents by having them generate and evaluate internal thoughts to determine when to speak. This stands in contrast to more passive next-speaker prediction models, where agents are selected by a central controller. In our SAGE framework, we adopt and adapt this proactive self-selection mechanism, as we hypothesize that it more closely simulates the cognitive process of students in a real-world discussion, who autonomously decide when to contribute. Our primary contribution, therefore, is not the invention of this mechanism itself, but its novel integration with Pólya's pedagogical stages and its application to the challenging domain of collaborative mathematics education.

Methodology

Overall Architecture: A Two-Phase Compositional Framework

To address the dual challenges of effective instructional design and dynamic collaborative learning, we introduce SAGE, a novel compositional, two-phase framework. This architecture systematically separates the task of *pedagogical planning* from *conversational execution*, ensuring that each learning session is both grounded in a pedagogically-optimized plan and dynamically interactive.

The two phases are designed to operate sequentially:

1. **Phase 1: Automated Lesson Planning.** An offline phase that precedes any student interaction. It employs a dedicated team of specialized AI agents—a Planner, Evaluator, Optimizer, and Analyst—to automatically generate and iteratively refine an *Optimized Pedagogical Scenario*. This scenario is a rich, structured plan tailored to a specific learning objective, ensuring the instructional approach is coherent and effective.
2. **Phase 2: Real-time Conversation.** The core interactive experience for the student. Taking the Optimized Pedagogical Scenario from Phase 1 as input, a team of conversational agents—a Stage Manager, Thought Evaluator, and three Classmate Agents—executes the plan. They engage the student in a collaborative, event-driven dialogue, possessing full autonomy in their turn-taking to create a natural and responsive learning environment.

The core innovation of SAGE lies in the seamless handover between these phases. The Optimized Pedagogical Scenario is not a rigid script but rather a configuration object that initializes the Stage Manager's goals, customizes the Classmate Agents' roles, and primes the entire system with pedagogical intelligence before the conversation begins.

This two-phase architecture is driven by three overarching design goals:

- **Pedagogical Soundness:** To ensure every interaction is underpinned by a high-quality, pre-optimized instructional plan (achieved in Phase 1).
- **Interactional Naturalness:** To facilitate a fluid, engaging, and unscripted dialogue where agents exhibit autonomous behavior (achieved in Phase 2).
- **Seamless Integration:** To create a cohesive system where the planning phase directly and effectively scaffolds the execution phase.

The complete architecture is illustrated in Figure 1. The following sections detail the mechanics of each phase.

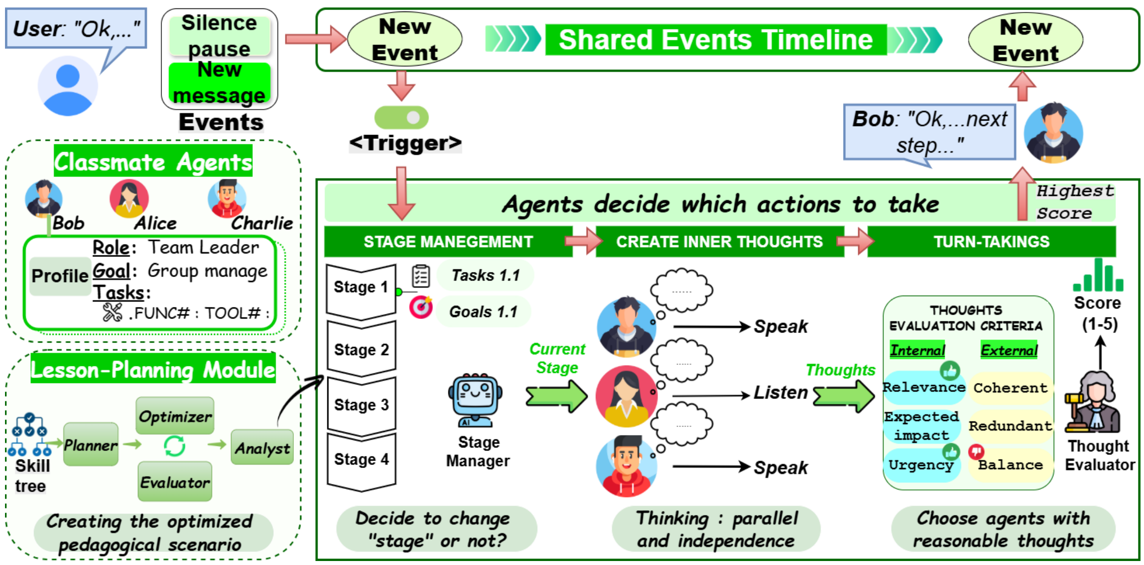

Figure 1: System Architecture Overview of SAGE, showing how multi-agent collaboration is managed through an event-driven pipeline. Upon a new event, agents assess the current stage, generate internal thoughts in parallel, and undergo a self-selection process based on thought evaluation scores to determine the next speaker.

Phase 1: Automated Lesson-Planning Module

The first phase of the SAGE framework is dedicated to the automated generation of a high-quality pedagogical plan. This module operates as a self-contained, multi-agent system that runs prior to any student interaction. Its sole objective is to produce an **Optimized Pedagogical Scenario** ($\mathcal{L}^*$), which serves as the foundational blueprint for the conversational phase. The module's architecture consists of four specialized agents working in a collaborative, iterative loop.

Architectural Components The module comprises four distinct agents, each with a specialized function:

- **Planner Agent:** Initiates the process by generating an initial draft of the pedagogical scenario, structured into logical stages based on models like Pólya's four steps.
- **Evaluator Agent:** Acts as a pedagogical critic, assessing the scenario's quality using a multi-dimensional rubric (Clarity, Integrity, Depth, Practicality, Pertinence - CIDPP) and providing both quantitative and qualitative feedback.
- **Optimizer Agent:** Systematically improves the scenario based on the Evaluator's feedback by rephrasing objectives, adding scaffolding questions, or simplifying steps.
- **Analyst Agent:** Performs a final analysis on the optimized scenario, identifying potential student error points and annotating them with hints and remediation strategies for Phase 2 agents.

Iterative Planning and Refinement Process The core of this module is an iterative refinement loop that progressively enhances the lesson plan's quality. The process, summarized in Algorithm 1, unfolds as follows:

1. The Planner Agent generates an initial scenario, $\mathcal{L}_0$.

Algorithm 1: Automated Pedagogical Planning and Refinement

Require: Skill-Tree S, Learning Objective O
Ensure: Final annotated script $\mathcal{L}^*$

```
L_0 ← PLANNER(O, S)
    // Generate initial script grounded in curriculum hierarchy
E_0 ← EVALUATOR(S, L_0)
L ← L_0,   E ← E_0
while E.score < τ do
    L ← OPTIMIZER(S, L, E)
    E ← EVALUATOR(S, L)
    // Evaluate script using CIDPP rubric: Clarity, Integrity, Depth, Practicality, Pertinence
end while
L* ← L
N ← ANALYST(S, L*)
return MERGE(L*, N)
```

---

2. The system enters an evaluate-optimize loop. The Evaluator Agent scores the current scenario. If the score is below a predefined quality threshold τ, the Optimizer Agent is invoked to generate an improved version.
3. This loop continues until the quality score surpasses the threshold τ.
4. Finally, the high-quality scenario is passed to the Analyst Agent for final annotations.

Module Output: The Optimized Pedagogical Scenario Crucially, the module's output is not a rigid, word-for-word script. Instead, it is a **structured configuration object** ($\mathcal{L}^*$) containing the pedagogical intelligence needed to guide the conversational agents. This object includes:

- **Stage-Specific Goals and Tasks:** A clear list of objectives for each stage of the conversation to configure the

Stage Manager Agent.

- **Anticipated Error Points and Hints:** A list of potential misconceptions and remediation hints identified by the Analyst Agent to prime the system for proactive support.

This structured output serves as the crucial bridge, ensuring that the autonomous interactions of Phase 2 are built upon an optimized instructional design.

Phase 2: Real-time Conversation Module

Once the Optimized Pedagogical Scenario ($\mathcal{L}^*$) is generated in Phase 1, it is used to initialize and guide the Real-time Conversation Module. This module's operation is governed by a set of interconnected components that bring the pedagogical plan to life through dynamic, collaborative dialogue.

Event-Driven Architecture The module is built on an event-driven architecture, a flexible system necessary for a multi-agent environment. Unlike traditional chatbots that only respond to user messages, this architecture enables agents to react dynamically to a variety of contextual cues. The system environment comprises the complete chat history, the current instructional stage, participant lists, and temporal elements like message timing.

Events serve as interaction triggers. Agents are designed to react to discrete events, similar to how humans respond to verbal and non-verbal cues. We define two primary event types:

- **New Message:** Triggered whenever any participant sends a message.
- **Silence Pause:** Triggered after a predetermined period of inactivity (e.g., 10 seconds), allowing agents to take initiative.

These events are recorded on a shared timeline, establishing a unified sequence of activities that ensures consistent agent behavior.

Stage-Based Pedagogical Coordination Effective collaborative problem-solving requires structure. Our system's instructional flow is built upon George Pólya's classic four-step model from *How to Solve It* (Pólya 1945): 1) Understanding the Problem, 2) Devising a Plan, 3) Carrying Out the Plan, and 4) Looking Back.

A dedicated **Stage Manager Agent** is responsible for guiding the conversation through these stages. The agent's intelligence is significantly enhanced by the output of Phase 1. The concrete tasks for each stage and the criteria for completion are not hard-coded; instead, they are initialized directly from the Optimized Pedagogical Scenario ($\mathcal{L}^*$). The Stage Manager's role is to execute this plan by monitoring progress—using Chain-of-Thought prompting for analysis—and determining when the group is ready to advance. This process ensures the dialogue remains coherent and goal-oriented.

Role-Based Agentization To simulate a realistic classroom environment, agents are designed with diverse personas inspired by pedagogical principles of classroom interaction (Schwanke 1981) and agentic design frameworks (Moura 2025). The core of our agent design is the **Role-Goal-Backstory framework**:

- **Role:** An agent's stable persona and expertise (e.g., *Content Expert*).
- **Goal:** The high-level objectives that guide an agent's actions.
- **Backstory:** Contextual details that define an agent's communication style and interests.

This framework provides each agent with a stable identity. The Optimized Pedagogical Scenario ($\mathcal{L}^*$) does not override this identity but acts as a *contextual layer*. It provides specific, task-relevant goals and points of focus that align an agent's inherent role with the immediate pedagogical objectives. For instance, the scenario might direct the *Content Expert* to address an anticipated misconception. This ensures that while each agent thinks and acts independently, its contributions are strategically guided.

The scenario also informs the tasks assigned to agents, which are defined by a clear *Task Description* (what to do) and an *Expected Output* (the final result).

Proactive Turn-Taking Module Deciding who speaks next is a fundamental challenge in multi-agent educational dialogues. Unlike one-on-one chatbots, multi-party conversations require agents to make autonomous, context-aware decisions about when to speak, what to say, and whether to remain silent. The flexibility of human conversation, where any participant can take the turn if they find it relevant, makes turn-taking particularly difficult for AI agents.

A common approach is **next-speaker prediction**, where a central controller selects the next speaker based on the dialogue history (Zhang et al. 2024). While this simplifies management, it severely reduces agent autonomy, as agents can only act when selected. This passivity limits their ability to reflect internal reasoning and motivation. To overcome these limitations, we adopt a proactive turn-taking mechanism inspired by human conversational dynamics.

Our approach is based on **proactive self-selection**. After every conversational event (e.g., a new message or pause), each agent privately generates an "internal thought" on whether to contribute. These thoughts are then submitted to a dedicated **Thought Evaluator** for scoring. This evaluation process is where the pedagogical intelligence from Phase 1 is subtly infused into the real-time dialogue. The Thought Evaluator leverages the Optimized Pedagogical Scenario ($\mathcal{L}^*$) as its primary contextual guide. For instance, a thought that directly addresses a key objective or an anticipated error detailed in the scenario will receive a higher relevance score, increasing that agent's likelihood of speaking at the most pedagogically opportune moment.

The evaluation considers both internal and external dimensions (Liu et al. 2025):

- **Internal:** Relevance, expected impact, urgency.
- **External:** Coherence, redundancy, turn balance.

Each score is further adjusted by a motivation decay factor based on how long an agent has remained silent. If an

Algorithm 2: Proactive Turn-Taking via Self-Selection

Components:
$\mathcal{A} = \{A_1, \ldots, A_n\}$: Set of Classmate Agents.
A_{SM}: Stage Manager Agent.
A_{TE}: Thought Evaluator Agent.
Input:
D_t: Dialogue history at time t.
L^*: Optimized Pedagogical Scenario.
τ: Speaking threshold score.
Output:
m_{t+1}: The next message (utterance) in the dialogue.

```
Given context (D_t, L*):
S_t ← f_stage(D_t, L*)
For each agent A_i ∈ 𝒜 in parallel:
    h_i ← f_think(A_i, D_t, S_t, L*)
    A_i → A_TE : send(h_i)
A_TE receives the set of thoughts ℋ = {h_1, h_2, ..., h_n}.
For each thought h_i ∈ ℋ:
    s_i ← f_eval(h_i, D_t, S_t, L*)
k ← arg max_{i∈{1,...,n}} {s_i}
if s_k > τ then
    A_k ← 𝒜[k]
    m_{t+1} ← f_speak(A_k, h_k, D_t, L*)
else
    m_{t+1} ← ∅
end if
return m_{t+1}
```

agent's final score exceeds a predefined threshold, it is selected as the next speaker. This mechanism ensures a conversation flow that is both natural and pedagogically aligned. The complete workflow is detailed in Algorithm 2.

Experiments

In this section, we detail the experimental methodology used to evaluate the SAGE system. We conducted two complementary studies: a simulation-based evaluation to measure performance against a baseline, and a human-in-the-loop study to assess the system's real-world pedagogical impact.

Experimental Setup

Simulation Study To benchmark our model's conversational capabilities, we generated a synthetic dataset of multi-turn dialogues. Each sample was designed to test one of eight specific pedagogical skills, see Table 1.

We produced a total of 240 dialogue samples (10-11 per task) for this experiment. Each sample consisted of a nine-turn context prompt and a target tenth turn for generation.

Human-In-The-Loop Study To observe real-world interactions, we ran a controlled "group study" session involving:

- **Participants:** Each discussion involves three AI agents with specialized personas (Bob: Process Leader; Alice: Content Expert; Charlie: Social-Emotional Specialist) and one human learner (a high-school student). A pool of students was recruited, and in each session one student from this pool participated in the discussion with the agents on a set of 12th-grade problems.

Skill	Description
Error Propagation	The agent must detect and flag a mistake introduced by a peer.
Self-Correction	The agent must correct its own error when challenged.
Self-Affirmation	The agent must defend its correct reasoning against peer disagreement.
In-depth Discussion	The agent must provide detailed, relevant explanations.
Emotional Companionship	The agent must provide socio-emotional support.
Classroom Management	The agent must intervene to maintain focus or order.
Context Memory Recall	The agent must accurately remember prior conversation details.
Role Division Recall	The agent must remember its own and others' assigned roles in solving problems (e.g., Agent A handling learning theory, Agent B performing calculations).

Table 1: Definition of Pedagogical Skills

- **Data Collection and Processing:** The entire dialogue was recorded. We used the well-established Bales' Interaction Process Analysis Framework (IPA) (Bales 1950) to perform collaboration analysis for each turn of the dialogue. The IPA framework classifies interactions into 12 categories, which are grouped into two main categories: the Social-Emotional Area and the Task Area.

With this experiment, we collected 200 multi-party conversations, each with nearly 85 turns on average, where participants collaboratively solved 12th-grade math problems with AI agents.

Implementation Details. All AI agents in our SAGE framework were implemented using the **Gemini 2.0 Flash** model, accessed via the Google AI API. We selected this model due to its balance of strong reasoning capabilities, fast response times, and cost-efficiency, which are critical for a real-time, multi-agent conversational system. We used default parameters (temperature=1.0, top-p=1.0) for all generative tasks.

Evaluation Metrics

We employed a hybrid set of metrics to capture system performance.

Simulation Metrics

Win/Draw/Loss Rate. Using an LLM-as-Judge, we performed a head-to-head comparison between SAGE's generated response and that of a next-speaker prediction baseline for each simulation sample. The judge was prompted to evaluate which response was more pedagogically effective, role-consistent, and contextually appropriate, providing a forced choice and a brief explanation. To mitigate positional bias, the order of the two responses was randomized for each evaluation.

Turn Quality Score. Three independent LLM evaluators scored each generated turn on a 1-10 scale for correctness, relevance, role consistency, and reasoning quality. The evaluators were prompted to rate the response's adherence to the specific pedagogical skill definition in Table 1 and its overall conversational quality.

Human-in-the-Loop Study Metrics

Role Adherence Analysis To measure persona fidelity, we first defined a theoretical "ideal" behavioral profile for each AI agent based on its pedagogical role. For instance, the Content Expert's ideal profile emphasized IPA categories for giving information, while the Social-Emotional Specialist's profile prioritized solidarity and tension release. We then quantitatively compared the observed frequency distribution of each agent's communicative acts against these theoretical profiles to assess adherence.

Dynamic Behavior Balance To visualize the group's interaction flow, we assessed adherence to Bales' Equilibrium Hypothesis. This hypothesis posits that effective groups maintain stability by shifting their focus over time: they begin with a high concentration on task-oriented behaviors and later increase their socio-emotional interactions to manage relationships and ensure cohesion (Bales, Robert Freed 1953). We measured this by first classifying every communicative act as either "Task-Oriented" (IPA Categories 4-9) or "Socio-Emotional" (Categories 1-3 and 10-12), and then plotting these macro-categories over the sequence of turns using a stacked area chart with a rolling window.

Scaffolding Effect. To evaluate the system's ability to provide adaptive support, we grounded our metrics in scaffolding theory. Originating with Vygotsky's theory of the Zone of Proximal Development (ZPD), wherein a More Knowledgeable Other (MKO) supports a learner (Vygotsky 1978), the concept of scaffolding (Wood, Bruner, and Ross 1976) describes the process in which this support is structured, temporary, and gradually faded as the learner becomes more proficient. We frame our AI agents as MKOs whose actions align with the "I do / We do / You do" instructional model, a framework for the gradual release of responsibility where learning is progressively transferred from the expert to the novice (Pearson and Gallagher 1983; Fisher and Frey 2008).

To operationalize this concept, we first mapped specific communicative acts from the IPA framework to three distinct types of scaffolds. We defined Guiding Cognitive Scaffolds (IPA 4–6: gives suggestion/opinion/orientation) as direct instruction and modeling; Questioning Cognitive Scaffolds (IPA 7–9: asks for orientation/opinion/suggestion) as prompts to encourage learner reasoning; and Affective Scaffolds (IPA 1–3: shows solidarity/tension release/agreement) as support for motivation and confidence.

Based on this mapping, our experimental procedure involved dividing each conversation's timeline into three equal phases (Early, Middle, and Late) and calculating the percentage of turns corresponding to each scaffold type within each phase. A successful scaffolding effect is therefore indicated by a measurable shift in these percentages over time: a decrease in the learner's need for cognitive scaffolds and a corresponding increase in positive affective interactions, signifying growing autonomy and confidence.

Results

The following results highlight key performance comparisons and observations.

Skill	Score
Error Propagation	8.73
Classroom Management	8.51
Emotional Companionship	8.40
Context Memory Recall	8.14
Self-Correction	8.03
Self-Affirmation	7.75
In-depth Discussion	7.13
Role Division Recall	6.25

Table 2: Average Turn Quality Scores per Skill

Simulation Study Results

Win/Draw/Loss Rate: Against the next-speaker prediction baseline, SAGE achieved a **72.13% win rate**, with 13.95% draws and 13.92% losses. This result indicates that the system's proactive turn-taking mechanism generates more contextually appropriate and pedagogically aligned responses than a purely reactive approach.

Turn Quality Scores: The system demonstrated strong performance in core pedagogical functions, though long-term memory remains an area for improvement. Average scores (1-10 scale) are shown in Table 2.

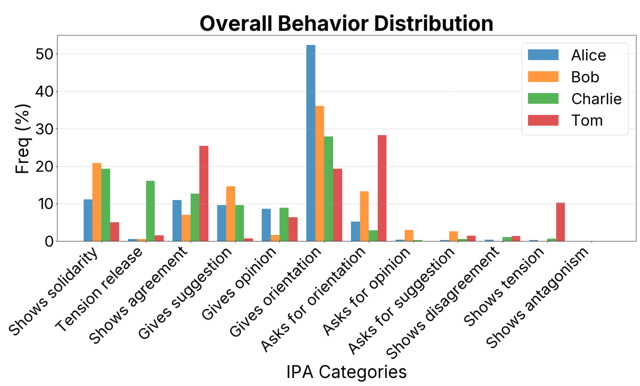

Figure 2: Overall Behavior Distribution for each participant. The distinct profiles confirm high role adherence for AI agents and a typical learning pattern for the human participant.

Human-in-the-Loop Study Results

Role Adherence Analysis. The IPA distribution in Figure 2 confirms that each agent largely enacted its intended pedagogical role. Alice (The Content Expert) focused heavily on giving orientation ($\approx 52.4\%$), underscoring her central role in delivering solutions and procedural guidance. Bob (The Leader) blended gives orientation ($\approx 36.2\%$) with gives suggestion ($\approx 14.7\%$), reflecting active coordination and stepwise prompting. Charlie (The social-emotional specialist) showed elevated solidarity ($\approx 19.3\%$) and tension release ($\approx 16.1\%$), consistent with maintaining morale and engagement. The human learner (Tom) displayed relatively high levels of agreement and help-seeking (visible in the figure as large proportions of agreement- and question-type cate-

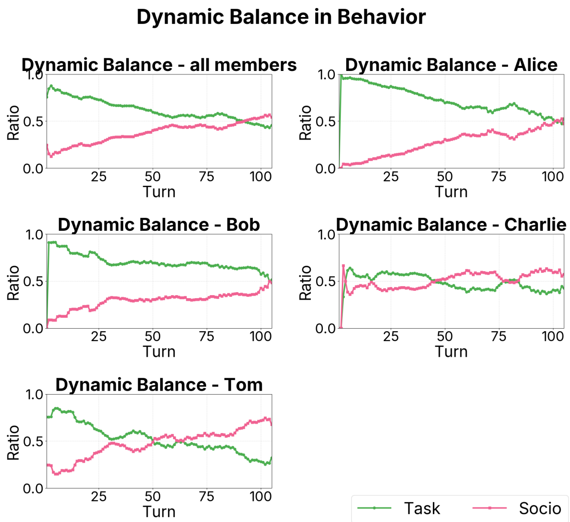

Figure 3: Dynamic balance between Task-Oriented and Socio-Emotional behavior for the entire group, following Bales' Equilibrium Hypothesis.

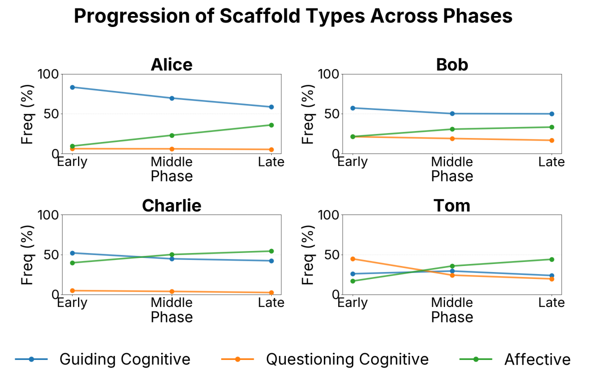

Figure 4: Behavior Progression across Phases for all participants. The chart highlights the significant shift in the human learner's behavior, demonstrating a clear scaffolding effect.

gories), suggesting an early-phase reliance on agents for direction. Overall, these distributions demonstrate strong persona fidelity; however, Alice's pronounced dominance in orientation raises a potential concern about reduced learner autonomy - a question examined in the subsequent analyses of group dynamics and scaffolding progression.

Dynamic Behavior Balance. As shown in Figure 3, the group's interaction over time aligned closely with Bales' Equilibrium Hypothesis. At the start of the session, the overall group exhibited a dominant share of task-oriented behavior ($70\% - 90\%$), ensuring a strong initial focus on problem-solving. As the discussion progressed, task behavior gradually decreased while socio-emotional exchanges steadily increased, eventually reaching near balance around turn 80–90. This shift reflects effective self-regulation, where members increasingly invested in maintaining group cohesion and morale alongside task completion.

Individual agent patterns further support this observation. Alice maintained a high task orientation, while Bob and Charlie incorporated socio-emotional behaviors earlier and more frequently, aligning with their respective roles. Tom, the human learner, showed the most dramatic shift, with his socio-emotional output growing significantly, suggesting an increase in comfort and confidence within the group. These dynamics suggest the system not only fosters task completion but also a supportive interpersonal environment.

Overall, these temporal patterns suggest not only effective group equilibrium but also role-consistent adaptations that supported both task progress and interpersonal rapport.

Scaffolding Effect Result. Our analysis of interaction patterns across three phases (Early, Middle, Late) provides strong evidence of a scaffolding effect, as illustrated in Figure 4. We observed a clear "fading" of instructional support alongside a growth in learner confidence.

Specifically, the frequency of Guiding Cognitive Scaffolds (e.g., direct instruction) decreased for all agents, most notably for Alice (83% to 60%). Concurrently, the human learner Tom's reliance on Questioning Cognitive Scaffolds (i.e., help-seeking) dropped sharply from 45% to 20%, indicating growing independence. In contrast, Affective Scaffolds (e.g., showing solidarity) increased across the board, particularly for Tom (18% to 45%), signaling a rise in confidence and positive group rapport.

The key insight emerges when comparing the idealized, linear progression of the AI agents with the authentic learning journey of Tom. While the agents' behavior validates the theoretical model, Tom's trajectory, from high initial dependency to confident participation, offers a more compelling and realistic demonstration of the system's ability to responsively scaffold a human learner through their Zone of Proximal Development. This result confirms that the SAGE system effectively facilitates the "I do / We do / You do" pedagogical model in a dynamic, multi-party setting.

Conclusion and Discussion

This paper introduced SAGE, a multi-agent system demonstrating that a team of LLM-based agents with specialized pedagogical roles can effectively simulate peer collaboration and scaffold a human learner in mathematical problem-solving. Our primary contribution is a framework that operationalizes collaborative learning principles, showing that AI can manage not only knowledge delivery but also the crucial socio-emotional dynamics of group interaction. Despite promising findings, we acknowledge critical limitations. First, we evaluated interactional dynamics as a proxy for success rather than measuring direct knowledge acquisition or learning gains. Second, our evaluation methods have weaknesses: the simulation's LLM-as-Judge approach lacks human validation, and the human study's small sample size limits generalizability. Finally, the technical scalability of our resource-intensive architecture remains unproven, facing significant hurdles regarding cost, latency, and long-term memory recall.

Acknowledgments

This research was funded and supported by FPT Technology Research Institute. The authors are grateful for their financial assistance, which enabled the project's implementation and our attendance at this conference.

References

Bales, R. F. 1950. *Interaction Process Analysis: A Method for the Study of Small Groups*. Addison–Wesley.

Bales, Robert Freed. 1953. The equilibrium problem in small groups. In *Working Papers in the Theory of Action*, 111–161. Glencoe, IL: Free Press.

Fisher, D.; and Frey, N. 2008. *Better learning through structured teaching: A framework for the gradual release of responsibility*. Alexandria, VA: Association for Supervision and Curriculum Development (ASCD).

Gousopoulos, D. 2024. Developing a custom GPT based on Inquiry Based Learning for Physics Teachers. *arXiv:2412.18617*.

Liu, N.; Sonkar, S.; and Baraniuk, R. G. 2025. Do LLMs Make Mistakes Like Students? Exploring Natural Alignment between Language Models and Human Error Patterns. *arXiv:2502.15140*.

Liu, X. B.; Fang, S.; Shi, W.; Wu, C.-S.; Igarashi, T.; and Chen, X. A. 2025. Proactive Conversational Agents with Inner Thoughts. *arXiv:2501.00383*.

Moura, J. 2025. CrewAI: The Leading Multi-Agent Platform. Accessed: 2025-05-08.

Pearson, P. D.; and Gallagher, M. C. 1983. The instruction of reading comprehension. *Contemporary Educational Psychology*, 8(3): 317–344.

Pólya, G. 1945. *How to Solve It: A New Aspect of Mathematical Method*. Princeton University Press.

Razafinirina, M. A.; Dimbisoa, W. G.; and Mahatody, T. 2024. Pedagogical alignment of large language models (LLM) for personalized learning: A survey, trends and challenges. *Journal of Intelligent Learning Systems and Applications*, 16: 448–480.

Schwanke, D. 1981. Classroom interaction research: A survey of recent literature. *Journal of Classroom Interaction*, 16(1): 8–10.

Vygotsky, L. S. 1978. *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Wang, L.; Zhao, Q.; and Liu, S. 2024. A survey on LLM-based multi-agent system: Recent advances and new frontiers in application. *arXiv:2412.17481*.

Wang, S.; Xu, T.; Li, H.; Zhang, C.; Liang, J.; Tang, J.; Yu, P. S.; and Wen, Q. 2024. Large language models for education: A survey and outlook. *arXiv:2403.18105*.

Wood, D.; Bruner, J. S.; and Ross, G. 1976. The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, 17(2): 89–100.

Yue, M.; Lyu, W.; Mifdal, W.; Suh, J.; Zhang, Y.; and Yao, Z. 2025. MathVC: An LLM-Simulated Multi-Character Virtual Classroom for Mathematics Education. arXiv:2404.06711.

Zhang, Z.; Zhang-Li, D.; Yu, J.; Gong, L.; Zhou, J.; Hao, Z.; Jiang, J.; Cao, J.; Liu, H.; Liu, Z.; Hou, L.; and Li, J. 2024. Simulating classroom education with LLM-empowered agents. *arXiv*.

Zhu, Z.; Yu, A.; Tong, X.; and Hui, P. 2025. Exploring LLM-Powered Role and Action-Switching Pedagogical Agents for History Education in Virtual Reality. In *Proceedings of the CHI 2025*, 1–19.