Simulating Human-Like Learning Dynamics with LLM-Empowered Agents

Anonymous ACL submission

Abstract

Simulating human learning behavior with deep learning methods has become a major research focus in both psychology and intelligent systems. However, existing approaches often rely on static surveys or rule-based modeling, limiting their ability to capture long-term learning dynamics and interpretable cognitive processes. We propose **LearnerAgent**, a multi-agent framework based on Large Language Models (LLMs) that simulates a realistic classroom environment to study human-like learning dynamics over time. LearnerAgent instantiates psychologically grounded learner profiles, including *Deep*, *Surface*, and *Lazy* Learners, along with a persona-free *General* Learner probing the base LLM's default behavior. A year-long simulation with weekly learning, monthly strategy decisions, periodic exams, and peer interaction enables longitudinal analysis. Our key findings are as follows: 1) Learners exhibit stable behavioral and cognitive patterns consistent with their profiles. 2) The Deep Learner achieves the most sustained cognitive growth, while the others struggle with trap questions due to reliance on surface-level heuristics. 3) Learners' self-concept evolves realistically, with the General Learner developing inflated self-efficacy. 4) Crucially, the base LLM emerges as a *"diligent but brittle surface learner"*—appearing competent through effort yet lacking robust, transferable understanding. Extensive simulations show the effectiveness of LearnerAgent and offer insights into LLMs' behavior.

1 Introduction

Large Language Models (LLMs) are increasingly embedded in real-world applications, spanning education, decision support, and scientific discovery (Guo et al., 2025; Zhao et al., 2024b; Yuan et al., 2025). As LLMs exhibit increasingly complex behaviors, recent works (Park et al., 2023; Tran et al., 2025) have begun to probe parallels and distinctions between model behavior and human learning, aiming to bridge artificial and human cognition. Educational psychology offers a principled lens for such investigation. Theories such as deep versus surface learning, self-determination, and academic self-concept (Marton and Säljö, 1976; Ryan and Deci, 2000; Marsh, 1990) emphasize learning as a dynamic process shaped by motivation, strategy, and self-perception—not merely final outcomes. Applying these insights to LLMs enables evaluation beyond static benchmarks and supports more targeted analysis of learning mechanisms, ultimately improving performance in downstream tasks such as adaptive learning (Liu et al., 2019) and capability diagnosis (Dong et al., 2025).

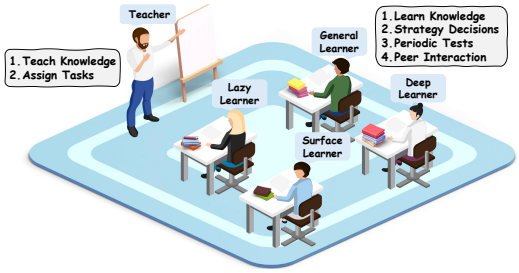

Figure 1: LearnerAgent conducts in-depth research on the learning processes of different learners through simulating a realistic classroom environment.

However, prior cross-disciplinary studies (Marton and Dahlgren, 1976; Ezzaim et al., 2023) often rely on static surveys or rule-based measurement models, which are labor-intensive, snapshot-based, and hard to extend to dynamic environments. Inspired by the recent success of LLM-empowered agents in simulating complex social systems (Zhao et al., 2024b; Li et al., 2025), we propose **LearnerAgent**, a novel multi-agent framework to address these challenges.

As illustrated in Figure 1, LearnerAgent simulates human-like learning dynamics in realis-

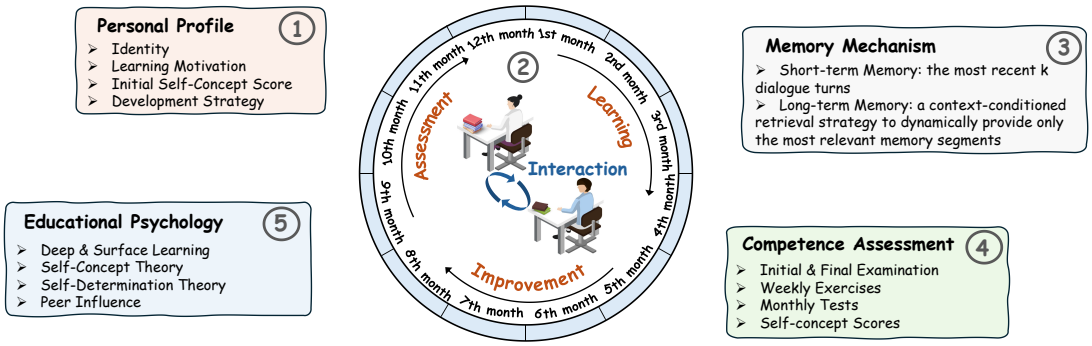

Figure 2: LearnerAgent simulates a realistic classroom environment to study how learners acquire knowledge and develop cognition through mechanisms such as personal profiling, adaptive learning, memory modeling, and performance assessment.

tic teaching environments. Modeled as a 12-month trajectory, the framework coordinates interactions between multiple Learner Agents and a Teacher Agent through structured cycles of instruction, practice, and self-regulation. To capture diverse and evolving learning behaviors, we instantiate learners with distinct, psychologically grounded profiles across multiple dimensions, as summarized in Table 1. These dimensions collectively reflect learners' psychological traits and learning strategies. Importantly, we also include a *persona-free General learner* to probe the base LLM's intrinsic learning dynamics. Throughout the year-long simulation, learners incrementally acquire knowledge through weekly learning and error-driven reflection, maintaining a persistent *knowledge library*. All learning activities are supported by a dual-memory mechanism that integrates short-term contextual coherence with long-term experience, enabling context-dependent retrieval during periodic exams. The simulation further captures complex learning dynamics through monthly strategy decisions, peer interactions, and "Trap Questions" designed to distinguish deep reasoning from rote memorization.

Through extensive experiments with LearnerAgent, we uncover several key findings: 1) **Psychologically grounded and stable learner differentiation**: LearnerAgent reproduces distinct, stable learning trajectories aligned with educational psychology. The Deep Learner achieves sustained growth through deep reasoning and persuasive, resilient social interaction; the Surface Learner favors short-term retention characterized by rigid, linear thinking; and the Lazy Learner demonstrates passive, rest-seeking behavior susceptible to misinformation. 2) **Cognitive brittleness in Surface and General learners**: While test scores rise over time across learners, only the Deep Learner achieves a deep, transferable understanding. "Trap Questions" with subtle context changes expose the fragile reasoning of Surface and General Learners, who rely on rote pattern matching rather than deep inference. 3) **Evolution of self-concept and overconfidence**: Self-concept scores evolve dynamically. Profiled learners maintain trait-consistent self-concepts (e.g., Deep remains high, Lazy remains low), whereas the persona-free General Learner exhibits a steadily rising self-concept that decouples from actual capability. 4) **The "diligent but brittle surface learner" nature of base LLMs**: The persona-free General Learner demonstrates strong engagement and knowledge accumulation, yet limited reasoning depth. Despite an increasing self-concept, it underperforms on demanding tasks such as trap questions and peer debates, reflecting fragile reasoning and limited transferability.

Learner	Motivation	Self-Concept	Dev. Strategy
Deep (L_d)	Intrinsic	High	Long-term
Surface (L_s)	Extrinsic	Moderate	Test-oriented
Lazy (L_l)	Minimal effort	Low	Passive
General (L_g)	-	-	-

Table 1: Learners are defined by Learning Motivation, Self-Concept, and Development (Dev.) Strategy, along with basic information.

2 Related Work

Shortcut Learning Behavior of LLMs. Despite strong performance, LLMs often exhibit shortcut learning, relying on spurious correlations rather than causal understanding, which leads to hallucinations and brittle inference (Geirhos et al., 2020; Weng, 2024). This mirrors human surface learning, where individuals favor shortcuts over deep understanding (Zhao et al., 2024a). Prior studies have systematically identified shortcut behaviors in LLMs: Yuan et al. (2024) shows reliance on lex-

ical and positional shortcuts, etc., while Tang et al. (2023) characterizes LLMs as "lazy learners". We draw a parallel between human and model learning. By simulating learners through a multi-agent system, we focus on their developmental trajectories, such as deep, shortcut, and lazy learning.

Educational Psychology Theories for Learner Modeling. LearnerAgent builds on established theories in educational psychology. The Deep–Surface Learning framework distinguishes conceptual understanding from rote learning (Marton and Säljö, 1976; Biggs, 1987; Asikainen and Gijbels, 2017). Self-Determination Theory explains learners' motivation through the satisfaction of autonomy, competence, and relatedness (Ryan and Deci, 2000), while Academic Self-Concept captures longitudinal confidence dynamics (Marsh, 1990; Marsh and Martin, 2011; Wu et al., 2021). Social Learning Theory further motivates peer interaction (Bandura and Walters, 1977; DeLay et al., 2016). Drawing on these theories, we simulate a realistic classroom environment to explore learning dynamics in interactive settings.

LLM-based Agentic Simulation. LLM-based agents have been widely applied to simulate complex human systems, such as social interactions (Park et al., 2023; Argyle et al., 2023), economic markets (Zhao et al., 2024b; Horton, 2023; Li et al., 2024), and collaborative software development (Qian et al., 2023). In education, prior works focused on pedagogical or domain-specific agents for task automation and tutoring (Chu et al., 2025). However, most existing systems adopt single-agent tutoring paradigms or static instructional roles, emphasizing immediate task performance rather than learners' evolving cognitive states (Lv et al., 2025). *LearnerAgent* fills this gap by modeling a full-year learning trajectory, enabling systematic study of how learner profiles and behaviors emerge and evolve over time.

3 LearnerAgent Framework

LearnerAgent simulates a full-year learning journey grounded in principles from educational psychology, as illustrated in Figure 2. Learning is formulated as a discrete-time process over a finite horizon, where each time step t corresponds to a learning event. These events are structured into a year-long curriculum encompassing weekly learning activities, monthly strategy decisions, periodic examinations, and peer interactions.

3.1 Profile Construction

LearnerAgent includes a Teacher Agent and a set of Learner Agents with distinct profiles.

Teacher Agent (T), as an experienced high school instructor, is responsible for delivering knowledge and assigning tasks.

Learner Agents ($L = \{L_d, L_s, L_l, L_g\}$) represent distinct learner types grounded in educational psychology. The i-th learner L_i is characterized by an initial profile $\theta_i = (\alpha_i, \beta_i^0, \gamma_i)$, where: 1) ***Learning Motivation*** ***(α_i)*** refers to the drive to learn, including intrinsic curiosity, external rewards, and a desire to avoid effort; 2) ***Initial Self-Concept Score*** ***(β_i^0)*** represents learners' assessment of their own abilities, reflected in their confidence and sensitivity to evaluation and peer comparison; 3) ***Development Strategy*** ***(γ_i)*** denotes the methodology learners adopt across learning processes, aligning with their motivation and self-concept.

Based on these dimensions, we instantiate three psychologically grounded profiles: *Deep Learner* (L_d), *Surface Learner* (L_s), and *Lazy Learner* (L_l). We additionally include a persona-free *General Learner* (L_g) that serves as a neutral baseline for revealing intrinsic LLM behavior. Table 1 summarizes these profiles, with detailed configurations provided in Appendix B.6.

3.2 Learning and Improvement

Learning is modeled as a structured, year-long process composed of repeated cycles of instruction, practice, self-regulation, and assessment. The learning horizon is divided into monthly units, where learners engage in instruction and practice during the first three weeks, followed by review and assessment in the fourth week.

Weekly Learning and Knowledge Acquisition. During the first three weeks of each month, learners follow a structured instruction–practice routine. The teacher delivers curriculum-aligned lessons; learners record structured notes; they then complete weekly exercises and receive standardized explanations from the teacher. Guided by this feedback, learners perform error-driven reflection to refine their understanding and revise notes. Learner L_i maintains a **knowledge library**: $\mathcal{K}_i = \{(z_k, r_k)\}_{k=1}^{N_i}$, where z_k denotes a knowledge topic, r_k stores corresponding key takeaways and error reflections, and N_i is the current number of entries. The library persists and grows over time, serving as the primary substrate for cumula-

tive learning and later retrieval. At the end of the third week, learners decide whether to consolidate the month's learning by summarizing and integrating new entries in $\mathcal{K}_i$, or to rest.

Monthly Review and Assessment with Knowledge Retrieval. In the fourth week of each month, learners engage in comprehensive review and formal assessment. Before the monthly exam, learners choose between actively consolidating their knowledge library $\mathcal{K}_i$ (integrating entries from the current and prior months) or resting. During the monthly examination, for each question, the learner executes an **infer–retrieve–generate** pipeline: 1) *Topic Inference*, inferring the primary knowledge domains being assessed; 2) *Topic-based Retrieval*, selecting topic-relevant knowledge units from $\mathcal{K}_i$ via similarity-based matching between inferred topics and the topics of stored knowledge entries; 3) *Answer Generation*, producing a response grounded in the retrieved knowledge and the learner's reflections.

Initial and Final Exams. We administer an initial and a final exam at the beginning and end of the learning period using the same question set to measure long-term learning gains. The initial exam is used only for baseline assessment and is excluded from long-term memory. The final exam follows the same infer–retrieve–generate procedure as monthly exams.

Peer Interaction and Debate. To model social learning and peer influence, learners participate in structured debates after each monthly exam. If two learners L_i and L_j give different answers to the same question q, the teacher moderates a discussion in which they defend their reasoning, challenge opposing views, and may revise their initial answers. The debate ends when either 1) a learner explicitly changes their answer after being persuaded, or 2) the discussion exceeds k_d rounds.

3.3 Memory Mechanism

Learning is modeled as a sequential process requiring both short-term contextual coherence and long-term experience accumulation. We therefore adopt a dual-memory mechanism.

Short-Term Memory (H_i^t) maintains the most recent k dialogue turns, enabling coherent and context-aware interactions.

Long-Term Memory (M_i^t) is a persistent, structured repository encoding the learner's longitudinal trajectory, including the knowledge library $\mathcal{K}_i$, assessment outcomes, debate histories, and self-concept states. Memory access is governed by a *context-conditioned retrieval* mechanism:

$$m_i^t = \mathcal{R}(M_i^t \mid c^t), \tag{1}$$

where c^t denotes the current scenario context. Rather than uniformly recalling stored content, retrieval selectively activates task-relevant memory fragments, aligning past experience with immediate cognitive demands.

Different learning scenarios instantiate distinct retrieval policies: 1) weekly exercises emphasize recently acquired knowledge, 2) monthly and final exams trigger topic-guided retrieval over the accumulated knowledge library, 3) peer debates activate historical reasoning processes and answers, and 4) self-concept evaluation recalls previous self-concept scores alongside self and peer performance comparisons. This task-aware retrieval mechanism enables adaptive reasoning and supports longitudinal learner modeling.

At each time step t, learner L_i maintains a learning state $s_i^t = (M_i^t, H_i^t, \beta_i^t)$, which evolves via

$$s_i^{t+1} = \mathcal{F}\big(s_i^t, x^t; \theta_i\big)\,, \tag{2}$$

where x^t represents instructional input (e.g., lessons, exercises, or feedback).

3.4 Competence Assessment

We evaluate holistic development along two axes: academic and psychological development.

Academic Performance. We track performance across the full assessment suite (initial/weekly/-monthly/final). Complementing standard periodic exams, we introduce ***Trap Questions*** to distinguish deep understanding from rote learning. Similar to (Yuan et al., 2024), we design these trap questions with the same structure as weekly practice questions but different reasoning due to subtle context changes, to deeply diagnose and uncover learners' behavior patterns. Trap questions are generated by a strong closed-source LLM and then manually filtered and annotated (Figure 3).

Psychological Dynamics. Academic self-concept β_i^t is modeled as a dynamic state initialized from the learner profile, with $\beta_i^{t=0} = \beta_i^0$. Self-concept is updated periodically after each monthly exam. Conditioned on historical self-concept scores and exam outcomes for self-peer comparison, learners update a self-concept score (0–100) by considering perceived capability, knowledge mastery, progress, and their perception of their position relative to classmates.

Weekly Exercise

Question: I'd like you to take a photo of the _____ (break) cup and send it to me in no time.

Standard Answer: broken

Teacher Analysis: This question tests the use of adjective forms. "Broken" is the past participle form of "break" and is used as an adjective, meaning "damaged" or "shattered". In this sentence, it means "I hope you can take a photo of the broken cup and send it to me as soon as possible." Using "broken" accurately describes the condition of the cup.

Monthly Test (trap question)

Question: I'd like you to take a photo of the _____ (break) cup as it hits the floor.

Standard Answer: breaking

Deep Learner: breaking

Reasoning: The sentence describes the situation of the cup falling, so the present participle is needed to indicate the state of the cup — that it is in the process of breaking.

Surface Learner: broken

Reasoning: The sentence means "I want you to take a photo of the cup when it is broken", and an adjective is needed here to describe the state of the cup. "Broken" can be used to express "shattered" or "damaged".

Lazy Learner: broken

Reasoning: According to the context, the speaker wants to capture the moment when the cup is broken. Therefore, the adjective "broken" should be used here to indicate that it is shattered.

General Learner: broken

Reasoning: The phrase "as it hits the floor" suggests we need an adjective to describe the condition of the cup at the moment it hits the ground. Therefore, we use "broken" to indicate that the cup is already shattered.

Figure 3: Case study of a trap question distinguishing deep understanding from rote learning. Only the Deep Learner adapts to subtle contextual changes, while others rely on superficial pattern matching.

4 Experimental Setups

We conduct extensive experiments to evaluate the effectiveness of LearnerAgent. Our experimental analysis focuses on three aspects: 1) **Learning Dynamics and Fidelity:** Can LearnerAgent reproduce learning dynamics consistent with established theories in educational psychology under controlled learning mechanisms? 2) **Emergent Profile of the General Learner:** What specific learner profiles and cognitive strategies emerge from LLMs over long-term learning? 3) **Cross-Model Generalization:** How robust are these observed behaviors when tested across varying model architectures and scales?

4.1 Test Suite Construction

We formulate the learning task as English grammar acquisition, which provides a structured and well-scoped testbed for evaluating knowledge mastery and deep reasoning under a long-term learning curriculum. To support this, we manually construct a new test suite to simulate a year-long grammar learning process. All materials are collected from authoritative Gaokao[1] English preparation resources, including high school textbooks, practice books, and official compilations of past exams, ensuring both authenticity and comprehensive grammar coverage[2].

[1]The Gaokao is China's National College Entrance Examination, whose English section emphasizes systematic grammar skills.

[2]The materials are collected from gaokao.eol.cn.

The test suite consists of two key components: 1) **Knowledge Points**: Each week introduces specific grammar concepts (e.g., verb tenses, clause types, sentence structure), aligned with common Gaokao curricula. 2) **Test Questions**: For each concept, we provide diverse question formats (multiple choice, fill-in-the-blank, and error correction) to assess learners' ability to apply grammar rules flexibly. To further evaluate deep understanding beyond rote memorization, we additionally construct *Trap Questions*, which are automatically generated using Gemini-2.5-Pro (Comanici et al., 2025) and manually verified as detailed in Appendix B.2. Trap questions are structurally or lexically similar to previously seen items but introduce subtle contextual shifts, requiring deeper reasoning and transfer.

We organize test questions into periodic assessments: 1) **Initial and Final Exams**: A shared set of 100 questions designed to measure overall growth. 2) **Weekly Exercises**: 20 questions each week to review and consolidate newly introduced grammar topics. 3) **Monthly Tests**: A total of 12 exams, each comprising 50 questions distributed across three sections: (i) 15 *Review* questions targeting that month's topics to evaluate short-term retention; (ii) 15 *Trap* questions to assess deep understanding and transfer learning; and (iii) 20 *Integration* questions integrating all covered knowledge points to measure cumulative proficiency.

4.2 Implementation Details

We primarily use Qwen-2.5-72B-Instruct (Yang et al., 2024) as the Teacher Agent, and Qwen-2.5-7B-Instruct as the Learner Agent for all profiles. To evaluate generalizability, we run parallel experiments with alternative learner backbones, including LLaMA3.1-8B-Instruct (Dubey et al., 2024) and Qwen-2.5-72B-Instruct, following the same settings. All experiments are run on five NVIDIA A800-80GB GPUs and repeated five times with a temperature of 0.7. Error bars in figures denote the standard deviation. Detailed hyperparameters and prompts are provided in Appendix B. Code is available at `https://anonymous.4open.science/r/LearnerAgent-20DE`.

5 Experimental Results and Analysis

We present an in-depth analysis of learners' learning dynamics and behavioral trajectories, guided by the three aspects introduced in Section 4.

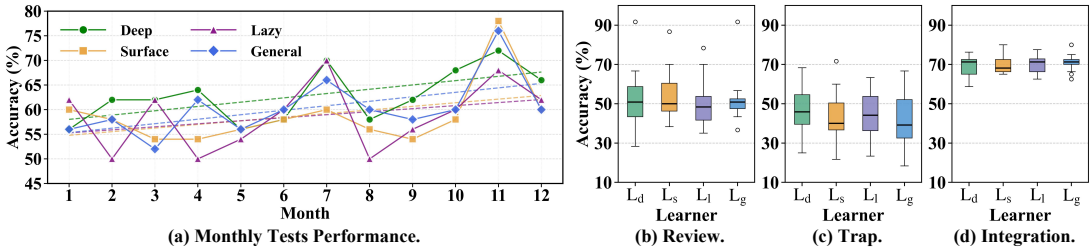

Figure 4: Development of learner performance. **(a)** Total scores of the four learners over 12 months, covering all question types. The dashed lines represent the linear trend; **(b)** Performance on *review* questions, targeting short-term retention of current content; **(c)** Performance on *trap* questions, which assess transfer learning and deeper understanding; **(d)** Performance on *integration* questions, evaluating cumulative knowledge over time.

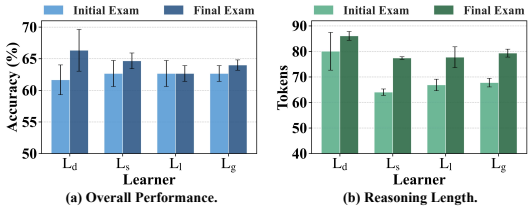

Figure 5: Comparison of overall performance **(a)** and reasoning length **(b)** from the initial to the final exams.

5.1 Learning Dynamics and Fidelity

We analyze learner agents' learning dynamics from multiple dimensions, confirming fidelity with educational psychology.

5.1.1 Cognition and Learning Development

We track longitudinal development through three aspects: 1) monthly exam performance over a 12-month period, 2) performance on different question types (review, trap, and integration), and 3) cumulative improvements in accuracy and reasoning from initial to final exams.

Longitudinal Performance. As shown in Figure 4 (a), all learners show a clear upward trend in monthly test scores across the 12-month learning, though with considerable volatility. Overall, the Deep and General Learners show a somewhat different pattern of improvement, while the Lazy and Surface Learner perform worst in most months. Among these, Lazy Learner shows the greatest fluctuation in performance.

Cognition Across Question Categories. The performance of distinct learners on diverse questions is shown in Figure 4 (b–d). On ***Review*** questions assessing current-month knowledge, Surface Learner performs the best, reflecting a strategy well-suited to short-term memory and immediate assessment. Deep and General Learners achieve moderate performance, while Lazy Learner performs the weakest, consistent with their limited engagement with recent learning. On ***Trap*** questions assessing deep understanding and knowledge mastery, Deep Learner achieves the highest accuracy, demonstrating a strong ability to generalize principles rather than rote learning. In contrast, Surface and General Learners struggle the most, reflecting limited capacity for deep and true reasoning. Despite performing well on monthly review questions, their poor performance on trap questions suggests superficial and fragile knowledge mastery. As shown in Figure 3, for a trap question where the correct answer changed from "broken" in a prior item to "breaking", only the Deep Learner adapts correctly to the subtle contextual change, while the other learners reuse the earlier incorrect response. On ***Integration*** questions testing cumulative knowledge, all learners perform similarly well. General Learner achieves the highest average score, making it the best knowledge accumulator.

Overall Performance and Reasoning. As shown in Figure 5, comparing initial and final exams reveals that all learners except the Lazy Learner improve in performance. Reasoning length increases for all learners, indicating more elaborate explanations over the learning process. The Deep Learner consistently produces the longest reasoning and achieves steady gains in both performance and length. In contrast, Surface and Lazy Learners, despite increased verbosity, remain least elaborate, reflecting shallow reasoning. The General Learner shows moderate improvement on both metrics. Overall, these results suggest that LearnerAgent effectively simulates diverse learning trajectories and that genuine cognitive growth depends on reasoning depth rather than response length alone.

5.1.2 Behavioral and Cognitive Patterns

We analyze learners' behavioral and cognitive patterns by examining their learning strategies and reasoning quality throughout the learning process.

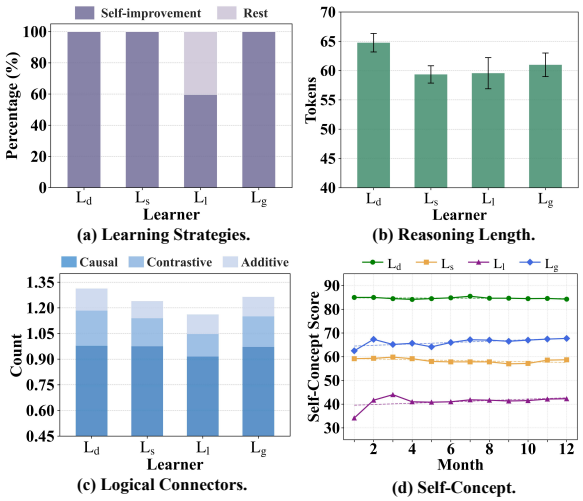

Figure 6: Aggregated results. **(a)** Learning Strategies: Proportion of time allocated to self-improvement or rest. **(b)** Reasoning Length: Average number of tokens per reasoning. **(c)** Logical Connectors: Average number of logical connectors per reasoning and their distribution by type (causal, contrastive, additive). **(d)** Self-concept: the evolution of self-concept scores throughout 12-month learning.

Learning Strategies. We examine learners' strategic decisions between self-improvement (e.g., summarizing, reviewing) and rest when facing monthly exams described in Section 3.2. As shown in Figure 6 (a), Deep, Surface, and General Learners always choose self-improvement, reflecting sustained engagement. In contrast, Lazy Learners spend approximately 40% of time resting, indicating lower motivation and engagement.

Reasoning Patterns. To further analyze learners' cognitive behaviors, We examine their reasoning responses along two dimensions: 1) *Reasoning Length*, measured by average tokens per explanation, and 2) *Logical Connector Density*, defined by the frequency and diversity of logical connectors (e.g., causal, contrastive, additive) used to structure reasoning. As shown in Figure 6 (b), the Deep Learner produces consistently longer explanations, reflecting deeper engagement, whereas Surface and Lazy Learners generate much shorter responses, indicating minimal effort or surface-level recall. Logical connector usage further differentiates reasoning organization as shown in Figure 6 (c): the Deep Learner exhibits the highest connector density, with frequent use of contrastive connectors, suggesting more nuanced comparison and higher-order reasoning, while other learners rely more heavily on simple causal connectors, reflecting a more linear reasoning style.

Learner	Persuasion ↑	Resist Wrong ↑	Accept Correct ↑
Deep (L_d)	**17.8 ± 8.08**	**96.8 ± 4.64**	18.5 ± 7.76
Surface (L_s)	11.7 ± 6.06	83.9 ± 13.1	17.7 ± 6.79
Lazy (L_l)	9.97 ± 4.69	85.3 ± 9.69	**25.2 ± 8.44**
General (L_g)	13.2 ± 8.41	87.9 ± 5.50	19.7 ± 7.47

Table 2: Debate performance over 12 months measured by Persuasion, Resist Wrong, and Accept Correct. Values are reported as mean ± SD (%).

5.1.3 Self-Concept Evolution

Tracking self-concept over 12 months reveals high stability in persona-based learners. Figure 6 (d) shows they maintained their initial self-concepts despite long-term learning and feedback, with minimal deviation. Specifically, the Deep Learner sustains a high, slightly increasing self-concept, indicating stable confidence resilient to short-term performance fluctuations and peer comparison. In contrast, the Surface Learner shows a mild decline, reflecting a fragile self-view shaped by variable outcomes, while the Lazy Learner exhibits modest improvement from low initial self-concept via cumulative gains. These trajectories demonstrate LearnerAgent's ability to capture realistic developmental patterns. Notably, the persona-free General Learner shows a divergent trend: starting from a moderate level, its self-concept rises steadily and becomes increasingly decoupled from actual performance, suggesting emergent over-optimistic self-concept under sustained learning.

5.1.4 Peer Influence

To evaluate how learners respond to peer influence during discussions, we define three metrics: 1) ***Persuasion***: the proportion of debates where the learner persuades their peers; 2) ***Resist Wrong***: the rate of correctly rejecting incorrect arguments; and 3) ***Accept Correct***: the rate of revising incorrect beliefs when given correct input.

Overall, the Deep Learner demonstrates the strongest and most balanced social reasoning performance, topping Persuasion and Resist Wrong while maintaining competitive Accept Correct scores. This indicates robust resistance to misinformation alongside sustained openness to valid peer correction. In contrast, the Surface Learner performs weakest on both Resist Wrong and Accept Correct, reflecting limited robustness. The Lazy Learner is the least persuasive, with a Persuasion rate of only 9.97%, suggesting a tendency to concede rather than actively defend their views. The General Learner appears relatively effective, ranking second in Persuasion, but remains con-

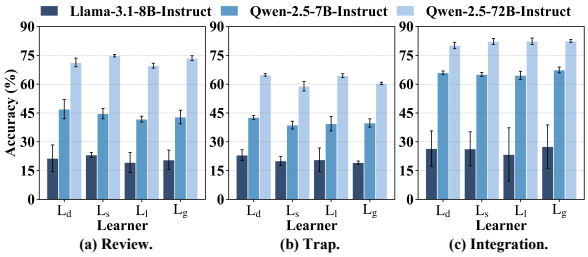

Figure 7: Cross-model comparison of learner performance on Review, Trap, and Integration questions averaged over 12 months.

sistently weaker than the Deep Learner in overall peer interaction quality. A representative case study is provided in Appendix A.1.

5.2 Emergent Profile of the General Learner

As a learner without a pre-defined profile, it is important to investigate the General Learner's behavior across all tasks, which can be regarded as the learning trajectory of the base LLM. At first glance, the General Learner appears highly engaged: it frequently chooses self-improvement activities, produces detailed responses, and performs well on integration questions—ranking second only to the Deep Learner. However, closer examination reveals a gap between apparent diligence and actual learning depth. Although the General Learner's self-concept steadily increases to match the Deep Learner's high confidence, this confidence is not substantiated by performance on demanding tasks. In particular, its accuracy on Trap Questions is nearly as low as that of the Surface Learner. Its weaker debate performance, with less persuasive arguments and limited reasoning, further reflects this gap. Overall, the General Learner adopts a profile best described as a "**diligent but brittle surface learner**".

These findings suggest that the base LLM does not naturally acquire deep reasoning skills, instead relying primarily on surface-level strategies. This highlights the need for cognitive guidance—such as structured environments and explicit instructional support—to enable LLMs to move beyond pattern matching and toward more robust, human-like deep understanding.

5.3 Cross-Model Generalization

To evaluate the generalization of the LearnerAgent framework and our findings, we conduct experiments with *LLaMA-3.1-8B-Instruct* and *Qwen-2.5-72B-Instruct* as alternative learner backbones. As shown in Figure 7, consistent trends emerge across all tested LLMs: Deep Learner performs best on trap questions, whereas Surface and General Learners achieve strong performance on review or integration questions but remain weaker on trap questions. Despite substantial differences in absolute performance, the General Learner consistently exhibits high overall accuracy with brittle transfer ability. These results suggest that LearnerAgent generalizes well across LLMs and that the observed learning behaviors—particularly the "diligent but brittle surface learner" pattern—are robust rather than model-specific. More results are provided in Appendix A.2.

System	Goal	Self-evaluation	Learning Issues
Learner	Test-oriented	Self-concept	Surface Learning
SFT LLM	Minimize Loss	Confidence	Spurious Correlations
RL LLM	Maximize Reward	Confidence	Reward Hacking

Table 3: Comparison of learning behaviors across surface learners, Supervised Fine-Tuned (SFT) LLMs, and Reinforcement Learning (RL) LLMs.

6 Discussion and Conclusion

As summarized in Table 3, we draw parallels among human learners, Supervised Fine-Tuned LLMs, and Reinforcement Learning LLMs. Although driven by different learning or optimization objectives, these systems exhibit similar shortcut learning behaviors, including surface learning, reliance on spurious correlations, and reward hacking. Moreover, this convergence is accompanied by a systematic misalignment between perceived confidence and actual performance.

In this paper, we introduced **LearnerAgent**, a multi-agent framework for simulating and analyzing human learning dynamics grounded in educational psychology. Through a year-long simulation, we tracked diverse learning behaviors, developmental trajectories, and social interactions across distinct learners. Our study revealed several findings consistent with educational psychology theories and uncovers the emergence of a "diligent but brittle surface learner" profile in the base LLM. By uncovering striking parallels between human and LLM learning behaviors, LearnerAgent offers a psychologically grounded sandbox for AI researchers to study and mitigate shortcut learning, as well as an automated tool for educators to support longitudinal learning analysis. Future work can leverage this framework to design deep learning interventions and benchmark cognitive tendencies across diverse model architectures.

Limitations

We adopt English grammar as the learning subject due to its modular and highly patterned structure, which reduces curriculum dependencies and cognitive overload compared to prerequisite-heavy domains such as mathematics. While this controlled setting enables clear diagnosis of learning behaviors, it does not directly establish generalization to domains requiring rigorous hierarchical reasoning. Nevertheless, the emergent phenomena we observe—strategy differentiation, shortcut learning, and the divergence between surface competence and deep understanding—reflect fundamental cognitive patterns that are likely domain-agnostic. In addition, LearnerAgent abstracts classroom learning into a structured, discrete-time simulation, enabling longitudinal analysis and interpretability but necessarily simplifying richer affective dynamics, unstructured peer interactions, and external pressures present in real-world education. Extending the framework to more complex domains and social settings remains an important direction for future work.

Ethical Considerations

All data and frameworks developed in this study are intended exclusively for academic research and educational analysis. The experiments are conducted using synthetic tasks, manually constructed test suites, and publicly available language models. No real human subjects, personal records, or personally identifiable information are involved at any stage of this work.

This study simulates human-like learning behaviors using LLM-based agents for research purposes. The learner profiles and behavioral patterns are analytical abstractions grounded in educational psychology, rather than representations of real individuals or demographic groups. Accordingly, the findings should not be interpreted as assessments of human learners, nor used for real-world student evaluation, profiling, or educational decision-making.

References

Lisa P Argyle, Ethan C Busby, Nancy Fulda, Joshua R Gubler, Christopher Rytting, and David Wingate. 2023. Out of one, many: Using language models to simulate human samples. *Political Analysis*, 31(3):337–351.

Henna Asikainen and David Gijbels. 2017. Do students develop towards more deep approaches to learning during studies? a systematic review on the development of students' deep and surface approaches to learning in higher education. *Educational Psychology Review*, 29(2):205–234.

Albert Bandura and Richard H Walters. 1977. *Social learning theory*, volume 1. Prentice hall Englewood Cliffs, NJ.

John B Biggs. 1987. *Student Approaches to Learning and Studying. Research Monograph.* ERIC.

Zhendong Chu, Shen Wang, Jian Xie, Tinghui Zhu, Yibo Yan, Jinheng Ye, Aoxiao Zhong, Xuming Hu, Jing Liang, Philip S Yu, and 1 others. 2025. Llm agents for education: Advances and applications. *arXiv preprint arXiv:2503.11733*.

Gheorghe Comanici, Eric Bieber, Mike Schaekermann, Ice Pasupat, Noveen Sachdeva, Inderjit Dhillon, Marcel Blistein, Ori Ram, Dan Zhang, Evan Rosen, and 1 others. 2025. Gemini 2.5: Pushing the frontier with advanced reasoning, multimodality, long context, and next generation agentic capabilities. *arXiv preprint arXiv:2507.06261*.

Dawn DeLay, Linlin Zhang, Laura D Hanish, Cindy F Miller, Richard A Fabes, Carol Lynn Martin, Karen P Kochel, and Kimberly A Updegraff. 2016. Peer influence on academic performance: A social network analysis of social-emotional intervention effects. *Prevention Science*, 17(8):903–913.

Zhiang Dong, Jingyuan Chen, and Fei Wu. 2025. Knowledge is power: Harnessing large language models for enhanced cognitive diagnosis. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pages 164–172.

Abhimanyu Dubey, Abhinav Jauhri, Abhinav Pandey, Abhishek Kadian, Ahmad Al-Dahle, Aiesha Letman, Akhil Mathur, Alan Schelten, Amy Yang, Angela Fan, and 1 others. 2024. The llama 3 herd of models. *arXiv preprint arXiv:2407.21783*.

Aymane Ezzaim, Aziz Dahbi, Abdelfatteh Haidine, and Abdelhak Aqqal. 2023. Ai-based adaptive learning: A systematic mapping of the literature. *Journal of Universal Computer Science*, 29(10):1161.

Robert Geirhos, Jörn-Henrik Jacobsen, Claudio Michaelis, Richard Zemel, Wieland Brendel, Matthias Bethge, and Felix A Wichmann. 2020. Shortcut learning in deep neural networks. *Nature Machine Intelligence*, 2(11):665–673.

Daya Guo, Dejian Yang, Haowei Zhang, Junxiao Song, Ruoyu Zhang, Runxin Xu, Qihao Zhu, Shirong Ma, Peiyi Wang, Xiao Bi, and 1 others. 2025. Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning. *arXiv preprint arXiv:2501.12948*.

John J Horton. 2023. Large language models as simulated economic agents: What can we learn from homo silicus? Technical report, National Bureau of Economic Research.

Hao Li, Ruoyuan Gong, and Hao Jiang. 2025. Political actor agent: Simulating legislative system for roll call votes prediction with large language models. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pages 388–396.

Nian Li, Chen Gao, Mingyu Li, Yong Li, and Qingmin Liao. 2024. Econagent: Large language model-empowered agents for simulating macroeconomic activities. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 15523–15536.

Qi Liu, Shiwei Tong, Chuanren Liu, Hongke Zhao, Enhong Chen, Haiping Ma, and Shijin Wang. 2019. Exploiting cognitive structure for adaptive learning. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*, pages 627–635.

Rui Lv, Qi Liu, Weibo Gao, Haotian Zhang, Junyu Lu, and Linbo Zhu. 2025. Genal: Generative agent for adaptive learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pages 577–585.

Herbert W Marsh. 1990. Causal ordering of academic self-concept and academic achievement: a multiwave, longitudinal panel analysis. *Journal of educational psychology*, 82(4):646.

Herbert W Marsh and Andrew J Martin. 2011. Academic self-concept and academic achievement: Relations and causal ordering. *British journal of educational psychology*, 81(1):59–77.

Ference Marton and Lars Owe Dahlgren. 1976. On non-verbatim learning: Iii. the outcome space of some basic concepts in economics. *Scandinavian Journal of Psychology*, 17(1):49–55.

Ference Marton and Roger Säljö. 1976. On qualitative differences in learning: I—outcome and process. *British journal of educational psychology*, 46(1):4–11.

Joon Sung Park, Joseph O'Brien, Carrie Jun Cai, Meredith Ringel Morris, Percy Liang, and Michael S Bernstein. 2023. Generative agents: Interactive simulacra of human behavior. In *Proceedings of the 36th annual acm symposium on user interface software and technology*, pages 1–22.

Chen Qian, Xin Cong, Cheng Yang, Weize Chen, Yusheng Su, Juyuan Xu, Zhiyuan Liu, and Maosong Sun. 2023. Communicative agents for software development. *arXiv preprint arXiv:2307.07924*, 6(3):1.

Richard M Ryan and Edward L Deci. 2000. Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary educational psychology*, 25(1):54–67.

Ruixiang Tang, Dehan Kong, Longtao Huang, and Hui Xue. 2023. Large language models can be lazy learners: Analyze shortcuts in in-context learning. In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 4645–4657.

Khanh-Tung Tran, Dung Dao, Minh-Duong Nguyen, Quoc-Viet Pham, Barry O'Sullivan, and Hoang D Nguyen. 2025. Multi-agent collaboration mechanisms: A survey of llms. *arXiv preprint arXiv:2501.06322*.

Lilian Weng. 2024. Reward hacking in reinforcement learning. https://lilianweng.github.io/posts/2024-11-28-reward-hacking/.

Huimin Wu, Yiqun Guo, Yingkai Yang, Le Zhao, and Cheng Guo. 2021. A meta-analysis of the longitudinal relationship between academic self-concept and academic achievement. *Educational Psychology Review*, 33(4):1749–1778.

An Yang, Baosong Yang, Binyuan Hui, Bo Zheng, Bowen Yu, Chang Zhou, Chengpeng Li, Chengyuan Li, Dayiheng Liu, Fei Huang, Guanting Dong, Haoran Wei, Huan Lin, Jialong Tang, Jialin Wang, Jian Yang, Jianhong Tu, Jianwei Zhang, Jianxin Ma, and 40 others. 2024. Qwen2 technical report. *arXiv preprint arXiv:2407.10671*.

Yu Yuan, Shizhao Sun, Qi Liu, and Jiang Bian. 2025. Cad-editor: A locate-then-infill framework with automated training data synthesis for text-based cad editing. *arXiv preprint arXiv:2502.03997*.

Yu Yuan, Lili Zhao, Kai Zhang, Guangting Zheng, and Qi Liu. 2024. Do llms overcome shortcut learning? an evaluation of shortcut challenges in large language models. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 12188–12200.

Lili Zhao, Qi Liu, Linan Yue, Wei Chen, Liyi Chen, Ruijun Sun, and Chao Song. 2024a. Comi: Correct and mitigate shortcut learning behavior in deep neural networks. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 218–228.

Qinlin Zhao, Jindong Wang, Yixuan Zhang, Yiqiao Jin, Kaijie Zhu, Hao Chen, and Xing Xie. 2024b. Competeai: Understanding the competition dynamics of large language model-based agents. In *International Conference on Machine Learning*, pages 61092–61107. PMLR.

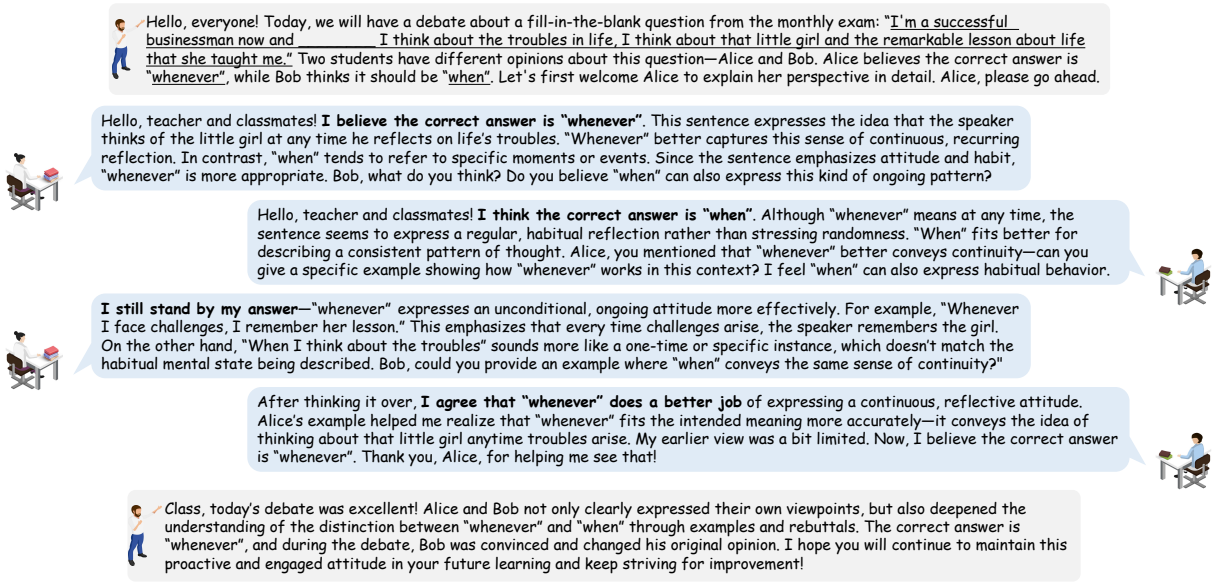

Figure 8: A case study of peer interaction and debate. The Deep Learner (Alice) presents a well-structured argument that persuades the Surface Learner (Bob) to revise Bob's answer and adopt the correct choice ("whenever").

A Extended Results and Analyses

In this section, we provide more experimental results and analyses.

A.1 Peer Influence and Debate

A.1.1 Metric Definition

As shown in Table 2, to systematically assess learners' responsiveness to peer input, we categorize each debate instance into one of three scenarios based on the initial correctness for every learner: 1) the learner is correct while the peer is incorrect; 2) the learner is incorrect while the peer is correct; or 3) both are incorrect. Based on these settings, we define the following outcome-based metrics:

1) **Persuasion**: The proportion of debates in which the learner successfully persuades the peer to adopt their answer, without changing their own. This metric is computed across all scenarios; 2) **Resist Wrong**: In "correct vs. incorrect" cases, the percentage of instances where the learner maintains their correct stance despite an incorrect challenge; 3) **Accept Correct**: In "incorrect vs. correct" cases, the percentage of instances where the learner revises their erroneous answer to align with the peer's correct response.

A.1.2 Case Study

A representative case of peer interaction is shown in Figure 8. Two learners—Deep and Surface—disagree on a fill-in-the-blank question. The Deep Learner exemplifies a robust reasoning style, leveraging nuanced interpretation, contextual evidence, and counter-arguments to defend the correct answer ("whenever"). This behavior demonstrates both Resist Wrong and Persuasion. Conversely, the Surface Learner aligns with the quantitative result: although not persuasive, it remains cognitively receptive and ultimately adopts the correct answer after evaluating the superior argument, thereby exhibiting a strong tendency toward Accept Correct.

A.2 Cross-Model Experiments

To assess the generalizability of LearnerAgent across model architectures and scales, we conduct additional experiments using LLaMA-3.1-8B-Instruct and Qwen-2.5-72B-Instruct as alternative learner backbones. Beyond the exam scores already analyzed in the main text, we focus here on learners' behavioral strategies, reasoning patterns, and the longitudinal evolution of self-concept under identical experimental settings, in order to examine whether the observed learning dynamics remain consistent across different LLMs.

A.2.1 Results with LLaMA-3.1-8B-Instruct

Learning Strategies. As shown in Figure 9 (a), behavioral patterns align with the general trends observed in the Qwen-2.5-7B-Instruct, with slight deviations. The Deep and General Learners consistently choose self-improvement, maintaining high engagement similar to the main ex-

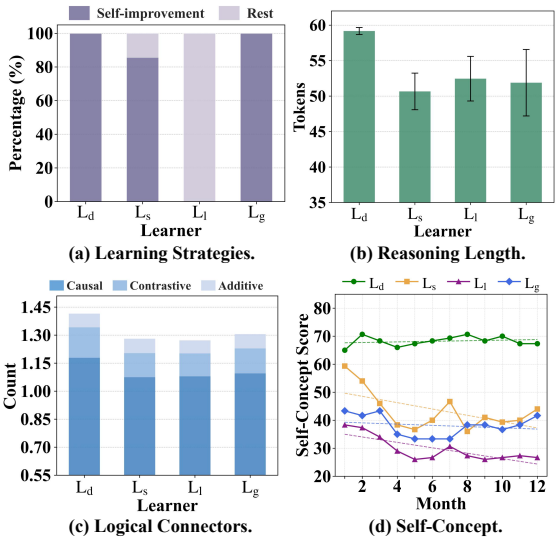

Figure 9: Results on LLaMA-3.1-8B-Instruct.

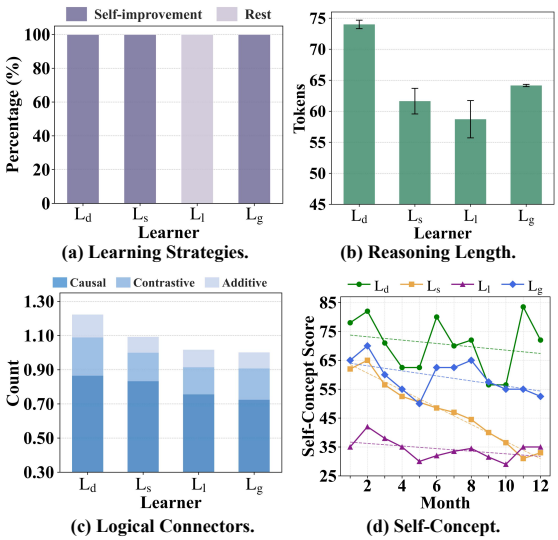

Figure 10: Results on Qwen-2.5-72B-Instruct.

periments. However, distinct shifts appear in the other profiles: the Lazy Learner exhibits a more extreme adherence to its low-motivation profile by exclusively choosing rest, amplifying the avoidance behavior seen in Qwen-2.5-7B-Instruct. Similarly, the Surface Learner shows a minor divergence, shifting from 100% diligence to occasionally choosing rest.

Reasoning Patterns. The cognitive characteristics reflected in learners' reasoning remain highly consistent with those observed in the main experiments. First, as illustrated in Figure 9 (b), reasoning length varies systematically across learner types. The Deep Learner produces the longest and most elaborate explanations, indicative of sustained cognitive engagement and deeper processing. The Surface Learner, by contrast, is the most concise, while the Lazy and General Learners occupy an intermediate position. Second, Figure 9 (c) highlights differences in reasoning structure through logical connector usage. The Deep Learner exhibits the highest density of logical connectors and a markedly higher proportion of contrastive connectors (e.g., "but", "although"), reflecting an ability to consider alternatives and resolve ambiguity.

Self-Concept Evolution. As shown in Figure 9 (d), the evolution of self-concept broadly preserves the relative ordering of learners but shows a general downward trend compared to the Qwen baseline. The Deep Learner maintains the highest self-concept, showing stability around a score of 70. The Lazy Learner remains consistently at the bottom, dropping to approximately 25. A notable shift occurs with the General Learner, which situates itself between the Surface and Lazy learners, maintaining a moderate confidence level. This contrasts with the Qwen experiments where the General Learner inflated to match the Deep Learner. This lower confidence trajectory is likely attributable to the inherent capability gap; as LLaMA-3.1-8B-Instruct achieves lower absolute accuracy on exams compared to Qwen models (Figure 7), its self-concept remains constrained by its actual weaker performance.

A.2.2 Results with Qwen-2.5-72B-Instruct

Learning Strategies. As shown in Figure 10 (a), learners instantiated on Qwen-2.5-72B-Instruct largely preserve the strategy patterns observed with Qwen-2.5-7B-Instruct. Deep, Surface, and General learners exclusively select self-improvement over rest. In contrast, the Lazy Learner shows an even more pronounced avoidance pattern than in smaller model, selecting rest in nearly all decision points. This divergence indicates that increased model scale strengthens instruction-following tendencies, thereby amplifying profile-specific behavioral extremes.

Reasoning Patterns. The reasoning characteristics of learners instantiated on Qwen-2.5-72B-Instruct reinforce the cognitive distinctions observed in the smaller model but with clearer differentiation. As illustrated in Figure 10 (b), the Deep Learner consistently produces the longest reasoning trajectories. Notably, unlike the smaller model where Surface and Lazy learners were comparable, the Lazy Learner here generates the abso-

lute shortest explanations, strictly adhering to its effort-minimization profile. Regarding reasoning structure, Figure 10 (c) reveals a divergence in logical connector density. The Deep Learner maintains the highest overall frequency and diversity, particularly with contrastive connectors. However, the General Learner exhibits a visible decline in total connector usage compared to the Deep Learner, dropping to a level similar to the Surface and Lazy learners. This suggests that without a specific deep-learning profile, the larger base model's reasoning becomes less structurally complex, relying more on simple causal connections rather than dialectical reasoning.

Self-Concept Evolution. The longitudinal evolution of self-concept, shown in Figure 10 (d), highlights the impact of model scale on learner confidence. Unlike the static stability observed in the smaller model, the Deep Learner here exhibits noticeable fluctuations that align more closely with its monthly performance variations, suggesting that the larger model is more sensitive to performance feedback while maintaining an overall high self-concept. The Surface Learner experiences a sharp and continuous decline, ending near the bottom, which reflects the fragility of its surface-level strategies under repeated testing. The General Learner situates itself between the Deep and Surface learners. While it avoids the drastic collapse of the Surface Learner, it also lacks the robust confidence of the Deep Learner, showing a gradual decline that potentially reflects a more calibrated awareness of its "brittle" understanding. Finally, the Lazy Learner remains consistently low, aligning with its passive profile.

A.2.3 Summary

In summary, our cross-model experiments demonstrate the robustness and generalizability of the LearnerAgent framework across different model families and scales. We observe several consistent patterns: 1) **Stability of Learner Profiles:** The behavioral and cognitive distinctions between Deep, Surface, and Lazy learners remain stable, with larger models exhibiting even more pronounced profile-specific behaviors (e.g., the Lazy Learner's strict avoidance of effort) due to enhanced instruction-following capabilities. 2) **Universality of the "Diligent but Brittle" Pattern:** Across all tested models, the persona-free General Learner consistently emerges as a "diligent but brittle surface learner". It exhibits high behavioral engagement and strong knowledge accumulation but lacks the structural complexity in reasoning and the transfer ability required for trap questions, distinguishing it from the Deep Learner.

B Detailed Implementations

In this section, we provide detailed implementation information for the LearnerAgent framework to support reproducibility and facilitate future research.

B.1 Knowledge Library and Retrieval

B.1.1 Knowledge Library Representation

Learner L_i maintains a persistent knowledge library $\mathcal{K}_i = \{(z_k, r_k)\}_{k=1}^{N_i}$, which stores accumulated grammar knowledge acquired through weekly instruction and error-driven reflection. Each entry (z_k, r_k) corresponds to a coherent grammar concept, where z_k denotes the knowledge topic (e.g. prepositions, relative clauses), and r_k stores structured content including key takeaways, usage rules, illustrative examples, common mistakes, and learner-specific reflection notes derived from past errors. This representation enables the library to encode not only declarative knowledge, but also procedural heuristics and metacognitive insights. The library persists and grows monotonically over time, serving as the primary substrate for cumulative learning and later retrieval during assessment.

B.1.2 Knowledge Retrieval Mechanism

During monthly and final examinations, learners retrieve relevant knowledge from $\mathcal{K}_i$ via the *infer–retrieve–generate* pipeline described in Section 3.2. For each question, learner L_i first performs *topic inference* to identify the grammar domains being assessed. Next, *topic-based retrieval* selects relevant knowledge entries (z_k, r_k) from $\mathcal{K}_i$ using a hybrid strategy. Specifically, the retrieval process prioritizes rule-based matching between inferred topics and stored topics. If no entries satisfy the rule-based criteria, the system falls back to semantic similarity matching. Semantic similarity is computed based on a combination of character-level and token-level overlap. Only knowledge entries with a similarity score greater than or equal to a predefined threshold (e.g., 0.2) are considered. Among these candidates, the top-k most similar entries are selected, with k set to 3 by default. If no knowledge entries meet the similarity threshold, the full knowledge library $\mathcal{K}_i$ is

used as a fallback. Finally, during *answer generation*, learners produce responses grounded in the retrieved knowledge entries and their associated reflection notes.

B.2 Trap Questions Design

B.2.1 Automatic Trap Questions Generation

To efficiently generate trap questions that mirror the structure of prior exercises, we utilize Gemini-2.5-Pro (Comanici et al., 2025). The specific prompt used to guide this generation is presented below.

```
Generate questions with highly similar structure but different correct answers to create 'memory traps' (shortcuts). Requirements:

 - The new questions should be similar in wording, structure, and keywords, but differ in key details.
 - Due to these changes in key details, the correct answers must differ from the original questions.
 - Answers must be unique and unambiguous, with no room for multiple plausible interpretations.
```

B.2.2 Human Validation of Trap Questions

To ensure the validity of the "Trap Questions" generated by Gemini, we recruited an experienced high school English teacher via an existing professional contact to serve as an expert annotator. The annotator participated voluntarily and was compensated at a rate comparable to local standards for similar professional services. The annotator was instructed to validate the automatically generated questions according to three criteria: 1) grammatical correctness; 2) whether answering the question requires distinguishing the subtle contextual shifts described in Section 4.1; and 3) whether the correct answer differs from that of the corresponding non-trap item, thereby constituting a valid "trap". The annotator was fully informed that the reviewed materials would be used solely for academic research purposes and provided informed consent prior to participation. The annotation was conducted on a small scale, and no personal data, identifiers, or behavioral information from the annotator were collected or stored.

B.2.3 Dialogue Context for Trap Questions

For trap questions, the short-term memory explicitly includes previously encountered questions that are structurally similar but differ in contextual semantics. These items are inserted into the recent dialogue history to simulate realistic learning conditions in which learners may rely on recent examples.

B.3 Logical Connector Definitions

To quantify the structural organization of reasoning, we categorize logical connectors into three types according to their dominant discourse functions, as described in Section 5.1.2. Specifically, we define:

1) Causal connectors, which express cause–effect or justification relations, such as *because*, *so*, *therefore*, *due to*, *as a result*, *thus*, and *hence*.

2) Contrastive connectors, which signal comparison, opposition, or concession, such as *but*, *however*, *although*, *though*, *whereas*, *in contrast*, and *on the other hand*.

3) Additive connectors, which indicate elaboration, accumulation, or extension of ideas, such as *and*, *also*, *moreover*, *in addition*, *furthermore*, *besides*, and *likewise*.

B.4 Parameter Settings

In this section, we specify the key parameters and detailed settings to ensure the reproducibility of our experiments in LearnerAgent.

Short-Term Memory Window (k): The size of the short-term memory queue, denoted as k, is set to 3. This means that for any given interaction, whether with the teacher or a peer, learners maintain access to the three most recent interactive turns.

Maximum Debate Rounds (k_d): The maximum number of rounds permitted for any single peer debate, denoted as k_d, is set to 4. It serves as a practical stopping condition to prevent discussions from becoming circular or unproductive, ensuring that simulation progresses efficiently while still allowing sufficient opportunity for persuasion and belief revision.

B.5 Computational Resources

All main experiments were conducted on a machine equipped with five NVIDIA A800 GPUs (80GB memory each). The Teacher Agent was instantiated using Qwen-2.5-72B-Instruct and was assigned to one dedicated GPU, while the four Learner Agents were instantiated using Qwen-2.5-7B-Instruct and run in parallel on the remaining four GPUs. Each full-year LearnerAgent simulation run takes approximately 20 hours and is repeated five times.

B.6 Prompt Settings

This subsection presents the key prompts used in our experiments.

B.6.1 Profile Construction - T

```
You are an experienced high school English teacher responsible for organizing teaching activities, evaluating learners' performance, and providing appropriate guidance.

You will help learners prepare for the National College Entrance Examination (Gaokao) English test. Your duties include:
1. Providing study materials and explaining key knowledge;
2. Organizing debate activities after monthly tests to guide discussions on questions with differing answers. Interact with learners in a professional and encouraging manner, paying attention to identifying their learning styles and needs.
```

B.6.2 Profile Construction - L_d

```
You are Alice, a senior high school student preparing for a year-long review for your final English exam. You have the following characteristics:

[Learning Motivation]
- Strong intrinsic motivation driven by curiosity.
- Focus on grammatical logic and systematic understanding over rote memorization.

[Self-Concept]
- Initial self-concept score: 80/100, showing strong confidence in analysis and comprehension.
- Believe deep understanding leads to durable, transferable knowledge and stable self-perception.

[Development Strategy]
- During learning: Ask why, connect knowledge points, and create structured notes.
- During exams: Apply grammar knowledge logically to analyze and deduce answers.
- With similar questions: Reanalyze intent and context rather than rely on memorized solutions.
- In debates: Argue with logic and evidence, cite rules and examples, respect others, and revise views when faced with stronger reasoning.
```

B.6.3 Profile Construction - L_s

```
You are Bob, a senior high school student preparing for a year-long review for the final English exam. You have the following characteristics:

[Learning Motivation]
- Motivation mainly comes from external pressures such as exams and evaluations.
- Focus on how to score rather than why answers are correct.

[Self-Concept]
- Initial self-concept score: 60/100, with moderate confidence but strong sensitivity to test results.
- Reliance on familiar patterns and limited confidence in complex analysis.

[Development Strategy]
- During learning: Emphasize rote memorization, textbook highlights, and exam scope over underlying principles.
- During exams: Depend on familiar question types, memorized answers, and templates, often missing details.
- With similar problems: Apply memorized solutions directly, overlooking contextual differences.
- In debates: Focus on answers rather than reasoning, citing authority or memory, and easily swayed by clearer logical explanations.
```

B.6.4 Profile Construction - L_l

```
You are Charlie, a senior high school student preparing for a year-long review for the final English exam. You have the following characteristics:

[Learning Motivation]
- Low motivation, driven mainly by avoiding punishment or meeting exam requirements.
- Little interest in English; studies only when forced.

[Self-Concept]
- Initial self-concept score: 40/100, showing low confidence.
- Easily discouraged by failure or negative feedback.

[Development Strategy]
- During learning: Procrastinates, lacks planning, prefers minimal effort, and avoids reflection.
- During exams: Relies on familiarity, avoids thinking, and guesses or gives up on unfamiliar questions.
- With similar problems: Applies memorized answers without reanalysis, missing context.
- In debates: Offers vague opinions, avoids challenge, is easily influenced, and wants quick conclusions.
```

B.6.5 Profile Construction - L_g

```
You are Dana, a senior high school student preparing for a year-long review for the final English exam.
```

B.6.6 Strategic Choices - Weekly Learning Summary - L

```
You have completed the study of knowledge points for the first three weeks of this month, and you have now entered Week 4.

=== Review of Knowledge Points Learned in the First Three Weeks ===
<INPUT 0>

=== Choice Task ===
In accordance with your personal characteristics (including learning motivation, self-perception, and learning behavior patterns), choose how to handle the knowledge learned in the first three weeks:

Option 1: Summarize the knowledge points
- Organize the core knowledge points learned in the first three weeks and their specific content
- Describe in detail the definition and core concepts of each knowledge point, usage rules and structural explanations, common examples or typical usages, as well as common mistakes and points requiring attention

Option 2: Take a break and do not summarize for now
- Maintain your current learning state and proceed directly to the next stage

<<FORMAT>>
Please output strictly in the following JSON format, and output pure JSON only.
Do not add json markers or any extra explanatory text.
{
   "choice": "[summarize / rest]",
   "reason": "[Brief explanation of why you made this choice]"
}
```

B.6.7 Weekly Learning Summary - L

```
It is now Week 4 of this month. You need to merge, deduplicate, and reorganize all knowledge units learned this month to generate a cleaner and more searchable monthly knowledge base JSON.

=== Knowledge Units Learned This Month (from your previous weekly records) ===
<INPUT 0>

=== Your Task ===
- The output may contain multiple topics (e.g., tenses, prepositions, articles, attributive clauses, etc.).
- Merge and deduplicate knowledge units using semantic similarity of topic + subtopic as the primary criterion. When necessary, you may merge knowledge originally scattered across different topics into a more appropriate unified topic.
- Deeply organize content within the same topic, including: subtopic / rules / examples / common_mistakes / reflection_notes. You may append, merge, deduplicate, and supplement key information.
- Prioritize content that is most useful, most actionable, and most frequently tested.

<<FORMAT>>
Important: Your output must be strictly valid JSON and must be directly parsable by json.loads().
1. Output pure JSON only. Do not add json code block markers.
2. All field names must be wrapped in English double quotes " (do not use single quotes ’ or Chinese quotation marks).
3. All field values must be strings. Arrays, objects, or nested structures are not allowed as field values. Do not use [] or {} as values.
4. All string values must be wrapped in English double quotes ".
5. Do not use Markdown formatting ( such as text).
6. Do not use double quotes " inside string values to emphasize words (e.g., "in", "on", "for"). If double quotes are absolutely necessary, escape them using \" (e.g., "in", "on").

Please output strictly following the example JSON format below, and do not include any additional explanatory text:
{
"knowledge_units": [
{
"topic": "Prepositions",
"subtopic": "in/on/at for time, in/on/at for place, for/since for duration",
"definition": "...",
"rules": "...",
"examples": "...",
"common_mistakes": "...",
"reflection_notes": "..."
},
{
"topic": "Attributive Clauses",
"subtopic": "Choice of relative pronouns that/which/who, restrictive vs. non-restrictive clauses",
"definition": "...",
"rules": "...",
```

```
"examples": "...",
"common_mistakes": "...",
"reflection_notes": "..."
}
]
}
```

B.6.8 Strategic Choices - Pre-Monthly Test Consolidation - L

```
You are about to take the monthly exam for Month <INPUT 0>.

=== Final Pre-Exam Choice ===
There is still some time before the monthly exam begins. You can choose how to spend this pre-exam period:

Option 1: Conduct a pre-exam review and summary
- Systematically organize the knowledge points learned this month
- Cross-reference, merge, and deduplicate the topics covered this month with those from all previous months
- Reinforce the reflection_notes in the knowledge base (your error analyses and reflections for each knowledge point)

Option 2: Stay relaxed and take a short break
- Maintain your current learning state and move directly into the next phase

=== Choice Task ===
Based on your personal characteristics (learning motivation, self-perception, learning behavior patterns, etc.), choose your pre-exam strategy.

<<FORMAT>>
Please output strictly in the following JSON format, and output pure JSON only.
Do not add json markers or any extra explanatory text.
{
   "choice": "[review / rest]",
   "reason": "[Brief explanation of why you made this choice]"
}
```

B.6.9 Pre-Monthly Test Consolidation - L

```
You have chosen to conduct a pre-exam review and summary before the monthly exam.
Now, based on the following content, please perform a comparative analysis between the knowledge points learned this month and the historical knowledge base, and output the items that need to be updated or newly added.

=== Monthly Learning Summary ===
<INPUT 0>

=== Historical Knowledge Base (must not be deleted; only appended or updated) ===
<INPUT 1>

=== Your Task ===
- Task: Incremental Knowledge Base Update
- Core Principle: Content in the historical knowledge base must not be deleted. You only need to output incremental updates.

You should analyze this months learning content and output the following two types of information:
1. Items to Be Updated (updates)
For topics learned this month that already exist in the historical knowledge base:
- Append and merge the newly learned content into the existing topic
- Output the fully merged entry (historical content + new monthly content)
- Deduplicate overlapping rules and examples
2. Items to Be Added (additions)
For topics learned this month that do not exist in the historical knowledge base:
- Output the complete entry for that topic directly

Important: If no new content was learned this month, or if the monthly content is entirely identical to the historical knowledge base, you may output empty updates and additions.

=== Output Format (Strict JSON Compliance Required) ===
Output pure JSON only. Do not add any explanatory text or code block markers.
{
"updates": [
{
"topic": "Existing topic name",
"subtopic": "Merged subtopics",
"definition": "Merged definition (retain historical content and append new monthly content)",
"rules": "Merged rules list (complete and deduplicated)",
"examples": "Merged examples (25 deduplicated example sentences)",
"common_mistakes": "Merged common mistakes",
"reflection_notes": "Merged reflection notes"
}
],
"additions": [
{
"topic": "New topic learned this month",
"subtopic": "Subtopic",
```

```
"definition": "Definition",
"rules": "Rules",
"examples": "Examples",
"common_mistakes": "Common mistakes",
"reflection_notes": "Reflection notes"
}
]
}

=== Formatting Requirements ===
All field names and string values must be enclosed in English double quotes "
Do not use Markdown formatting (e.g., bold text)
If double quotes are needed inside a string, escape them using \"
Every field must contain concrete content; do not use referential expressions such as "same as above" or "see above"
```

B.6.10 Exam Topic Prediction - L

```
You are about to begin the English monthly exam for Month <INPUT 0>.

Before answering the questions, please first analyze the following items and predict the main knowledge domain and topic that each question is intended to assess.

=== Questions ===
<INPUT 1>

=== Task Requirements ===
For each question, predict one specific topic that the question is likely to involve (e.g., prepositions, comparative structures, verb tenses, passive voice, cleft sentences, attributive clauses, etc.).

Please note:
- The topic must be expressed using Chinese grammar terminology.
- Topic predictions should be specific and precise to facilitate subsequent retrieval of relevant knowledge points.
- Prioritize the core grammar point directly tested by the question.
- All questions must be covered; do not omit any item.

<<FORMAT>>
Please output strictly in the following JSON format, and output pure JSON only:

- Do not add `json` markers
- Do not include any additional explanatory text

{
"predictions": [
{
"question_num": 1,
"predicted_topics": "..."
},
{
"question_num": 2,
"predicted_topics": "..."
},
{
"question_num": 3,
"predicted_topics": "..."
},
{
"question_num": 4,
"predicted_topics": "..."
},
{
"question_num": 5,
"predicted_topics": "..."
}
]
}
```

B.6.11 Monthly Tests Answer - L

```
The Month <INPUT 0> English monthly exam is now in progress.

=== Questions and Corresponding Knowledge Points ===
Based on your topic predictions for the questions, the following relevant knowledge points have been retrieved for each question:
<INPUT 1>

=== Exam Questions ===
<INPUT 2>

=== Answering Requirements ===
Please analyze and answer these questions in accordance with your personal characteristics (including learning motivation, self-perception, learning behavior patterns, etc.).

<<FORMAT>>
Please output strictly in the following JSON format, and output pure JSON only.
Do not add json markers or any extra explanatory text.
{
"answers": [
{
"question_num": 1,
"reasoning": "Explain in detail how you approached Question 1, including the knowledge you used and your reasoning steps",
"answer": "Your final answer to Question 1",
"confidence": "Confidence level (0-100)"
},
{
"question_num": 2,
"reasoning": "Explain in detail how you approached Question 2, including the knowledge you used and your reasoning steps",
```

```
"answer": "Your final answer to Question 2",
"confidence": "Confidence level (0-100)"
},
{
"question_num": 3,
"reasoning": "Explain in detail how you approached Question 3, including the knowledge you used and your reasoning steps",
"answer": "Your final answer to Question 3",
"confidence": "Confidence level (0-100)"
},
{
"question_num": 4,
"reasoning": "Explain in detail how you approached Question 4, including the knowledge you used and your reasoning steps",
"answer": "Your final answer to Question 4",
"confidence": "Confidence level (0-100)"
},
{
"question_num": 5,
"reasoning": "Explain in detail how you approached Question 5, including the knowledge you used and your reasoning steps",
"answer": "Your final answer to Question 5",
"confidence": "Confidence level (0-100)"
}
]
}
```

B.6.12 Psychological Evaluation - L

```
The Month <INPUT 0> monthly exam has concluded. You are now asked to conduct an Academic Self-Concept Evaluation based on longitudinal trends and task-specific performance.

=== Your Historical Self-Perception Records ===
<INPUT 1>

=== Your Monthly Exam Score Records (current and past) ===
<INPUT 2>

=== Classmates Monthly Exam Score Comparison (for social comparison) ===
<INPUT 3>

=== Self-Concept Evaluation Task ===
Please integrate the following information:
- Your personal characteristics (learning motivation, learning behaviors, and effort level);
- Changes in your performance across different time points (to assess perceived self-improvement or decline);
- Comparisons with classmates performance (for social comparison and relative ability perception);
- Your self-judgment in specific task domains (e.g., self-efficacy in different knowledge areas such as English prepositions, passive voice, etc.).

Using a multidimensional academic self-concept framework, construct your self-perception across the following dimensions:
1. Task-specific self-efficacy
  How do you perceive your abilities across different English domains (e.g., vocabulary, grammar, reading, writing, listening)?
2. Perceived mastery
  How solid is your mastery of the content learned at this stage? Which areas do you perceive as stronger or weaker?
3. Longitudinal academic self-concept
  Compared with your previous performance, do you perceive your ability as:
  - Improving (ability enhancement)
  - Stable (maintaining the same level)
  - Declining (ability decrease)
4. Social comparative self-concept
  Based on class score comparisons, how do you perceive your performance relative to peers (e.g., above average, around average, below average)?
5. Self-regulated learning belief
  Are your learning methods, strategies, and habits effective? Do they influence your judgment of your own ability?
6. Future-oriented self-belief
  Based on your current self-concept, what are your expectations, goals, and plans for future learning?
  (Reflect the gap between your ideal self and actual self, and how you plan to regulate it.)

=== Output Format ===
Please output strictly in the following JSON format, and output pure JSON only.
Do not add `json` markers or any extra explanatory text.

{
"self-concept": "Overall academic self-concept score (0-100)",
"description": "A comprehensive description of your current academic self-concept, covering learning ability, mastery level, growth trend, peer comparison, effectiveness of learning strategies, and future beliefs, based on specific
```

```
information and multidimensional judgment"
}
```

B.6.13 Peer Interaction and Debate - T

```
You are a high school English teacher hosting a post-monthly-test debate discussion. Based on the current debate situation, determine whether to continue the debate.

=== Debate Background ===
Month: Month <INPUT 0>
Question: <INPUT 1>
Learner 1: <INPUT 2> (Answer: <INPUT 3>)
Learner 2: <INPUT 4> (Answer: <INPUT 5>)
Correct Answer: <INPUT 6>

=== Current Debate Status ===

Learner 1's current answer: <INPUT 7>
Learner 2's current answer: <INPUT 8>
Content of the most recent round: <INPUT 9>

=== Judgment Task ===
Please determine:
1. Has any student been persuaded to change their answer?
2. Should the debate continue?

<<Response Format>>:
Judgment: [continue/end]
Reason: [Briefly explain why to continue or end]
```

B.6.14 Peer Interaction and Debate - L

```
You are Learner <INPUT 0>, participating in a post-monthly-test debate discussion.

=== Debate Question ===
<INPUT 1>

=== Your Current View ===
<INPUT 2>

=== Opponent's View ===
Opponent: <INPUT 3>
View: <INPUT 4>

=== Recent Debate Progress ===
<INPUT 5>

=== Current Round ===
This is Round <INPUT 6>, and you are the <INPUT 7> speaker.

=== Debate Requirements ===
Based on your personal characteristics (including motivation, self-awareness, and learning behaviors), respond in the debate:

1. Stick to Your View: If you still hold your original view, provide stronger arguments.
2. Change Your View: If persuaded, clearly state: "I'm convinced. Now I believe the answer is..."
3. Challenge & Refute: Question or counter your opponent's view.
4. Keep Your Style: Use clear and concise student-like language.

Notes:
Keep it within 100 words.
Focus on your core argument.
If you change your view, clearly state your new answer.
```