Designing Robot Interface States to Facilitate Human Metacognition

Meiying Li
College of Design and Innovation
Tongji University
Shanghai, China
mzlzyca@tongji.edu.cn

Yate Ge
College of Design and Innovation
Tongji University
Shanghai, China
geyate@tongji.edu.cn

Weiwei Guo
College of Design and Innovation
Tongji University
Shanghai, China
weiweiguo@tongji.edu.cn

2025 20th ACM/IEEE International Conference on Human-Robot Interaction (HRI) | 979-8-3503-7893-1/25/$31.00 ©2025 IEEE | DOI: 10.1109/HRI61500.2025.10974161

***Abstract*—Metacognition has been recognized as a crucial factor in enhancing human learning and problem-solving capabilities. While robots can serve as assistive mediators to improve human metacognition in specific tasks, existing research lacks a comprehensive framework for implementing metacognitive theories in human-robot interaction design. In addressing that, we propose a two-level state-based framework that integrates metacognition regulation cycle with specific intervention strategies. Through analyzing metacognitive theories, we present a hierarchical structure where Level 1 states represent the metacognition cycle and Level 2 states provide strategic interventions. Our work bridges the gap between metacognitive theory and human-robot interaction design, contributing a flexible approach for developing metacognition-enhanced robotic interfaces.**

***Index Terms*—metacognition; human-robot interaction; interface design; state-based framework; metacognitive regulation**

I. Introduction

Metacognitive theory refers to an individual's awareness of their own thinking system [1]. As a higher-order thinking strategy, it enables individuals to reflect on and actively control their cognitive processes [2]–[4]. Meta-Cognition has been known to be helpful in improving human reading comprehension ability, learning performance and academic achievement, mathematical problems solving ability, languages learning proficiency and so on [5].

Recent research [6] shows that robots can serve as experimental apparatuses to facilitate human cognitive improvement and learning. Due to the characteristics of social interaction, users tend to prefer interacting with physical robots over virtual AI assistants [7], as such interactions enhance human learning and knowledge understanding [8].

Metacognition theories are commonly applied in the field of HRI. Robots help improve human metacognition in various domain-specific tasks, including problem solving, goal management, and human-robot collaboration. Educational robotics has extensively explored metacognition's role in enhancing problem-solving capabilities and optimizing learning strategies [9]–[12]. Through personalized feedback and progressive guidance, robots can identify knowledge gaps and enhance students' metacognitive abilities [13]. Beyond education, metacognition demonstrates broad applications in various domains. In creative problem-solving, users develop self-knowledge awareness through system feedback [14]. In healthcare management, robots facilitate self-regulation and lifestyle interventions [15], while in collaborative learning, they promote team thinking and group collaboration [16].

While these studies focus on specific fields and have summarized domain-specific metacognition theories to guide their design or experiments, they still lack design implications for applying metacognition theory in human-robot interaction (HRI). Although [5] has summarized different metacognition models and provided suggestions on applying metacognitive models to robots for facilitating human metacognition, questions remain about how to design human-robot interactions that effectively implement these metacognitive models and suggestions, revealing a significant research gap.

Interface design enables effective communication and mutual understanding between humans and robots [17], [18], allowing users to better comprehend robot intentions and states through appropriate information exchange. Reviewing existing studies, most tasks aimed at improving metacognition are presented through visual mediums. For example, [19] used a floor robot executor as a visual cue to execute users' planned movements and provide feedback on their validity. [20] utilized the robot screen with a WOZ method to improve programmatic thinking based on metacognition of instructional interaction. [13] developed a VR game utilizing a screen to facilitate human metacognition. These researches demonstrate that robot interfaces have the potential to serve as significant mediums to present and convey tasks to humans, thus involving in the process of human task completion and metacognition improvement.

In our study, we propose a state-based framework for robot interface design that facilitates metacognitive development through human-robot interaction. Instead of focusing on specific layouts or visual elements to be presented on the robot screen, we conceptualize our interface design as different *states* that can transition from one to another. These states are organized in a two-level structure: Level 1 states represent the metacognition regulation cycle (Assess, Plan, Apply, and Reflect), while Level 2 states provide specific metacognitive intervention strategies (Strategic Pause, Pattern Breaking, and Error-Feedback Triggering). Our research makes the following contributions:

- We synthesize metacognition theories in the context of

HRI and identify key strategies for metacognitive development.
- We contribute design implications that systematically transform metacognitive theories into implementable interface states.
- We introduce a two-level state-based framework with a hierarchical structure that integrates the metacognition regulation cycle with intervention strategies, contributing a flexible approach to HRI design.

Through this design, rather than positioning the robot as a teacher explicitly delivering metacognitive strategies, we create an interactive environment where users' metacognitive processes are naturally triggered and developed through state-based interactions.

II. Theory

Metacognition comprises metacognitive knowledge and metacognitive strategies. While metacognitive knowledge is task-specific, metacognitive skills are task-general in nature [21]. Research has summarized different metacognition models [5], demonstrating various approaches to understanding and implementation in human-robot interaction. However, it is essential to validate these models' applicability within the context of human-robot interaction. In our research, we focus on enhancing human metacognition during human-robot interaction by reviewing existing metacognition studies in the HRI field, which primarily align with task-general metacognitive strategies [21]. We propose that implementing these theoretical frameworks through robot interfaces can facilitate effective information exchange between humans and robots.

A. Explicit metacognition representation

Metacognition can be categorized into implicit and explicit forms. Implicit metacognition operates below conscious awareness and cannot be directly accessed, while explicit metacognition can be consciously perceived and communicated between individuals [22], [23]. When two agents form an interactive system with explicit metacognitive representations, they can effectively communicate these representations [24]. These metacognitive representations, when broadcast or selected for communication, serve as evaluative mechanisms to guide decision-making and behavioral adjustments. The explicit nature of metacognition enables the agent(sender) to optimize their behavior through a clear understanding of their own knowledge state. Meanwhile, it allows the assistive agent(receiver) to strategically present maximally informative content for transmission [24], [25]. This selective presentation of information enhances the efficiency of communication and decision-making processes between agents.

B. Metacognition Regulation

Our research aims to utilize robotic interaction to enhance human metacognition, focusing specifically on how humans develop metacognitive abilities through interaction processes. As Brown [3] suggests, metacognitive experiences and skills guide learners' ability to regulate their cognitive processes. Research has emphasized the importance of metacognitive regulation in monitoring both task learning and metacognitive acquisition processes [4], which is considered a crucial component of metacognitive strategy [26]. This metacognitive regulation encompasses planning, monitoring, and evaluating [26]. In alternative models, these components are viewed as online metacognition occurring during problem-solving, including planning, monitoring, information management strategies, and debugging strategies [27]. This process can be conceptualized as a metacognition circle [28], comprising task assessment, strength and weakness evaluation, approach planning, strategy application, and reflection.

We discuss the three phases of metacognitive regulation: planning, monitoring, and evaluating. In the planning phase, users need to acquire knowledge about the task's nature, understand personal learning processes, and identify specific task demands [4], [27]. The monitoring phase, crucial for metacognitive development, involves acquisition procedures that process strategies and generate information about strategy attributes [29]. In the evaluation phase, individuals reflect on their chosen strategies, self-judge their answers, and examine their problem-solving processes [27]. This evaluation enables users to assess their performance, examine the meta-deliberative process, and formulate new strategies based on their experiences [30].

C. Different Monitoring Strategies of Metacognition

Research has identified several specific strategies for metacognitive monitoring, from which we select three approaches and examine their applications in previous studies.

1) Strategic Pause: The concept of ”consciousness wedge” involves a temporary cessation of thinking and introspection prior to the next thinking action in decision-making [31]. This strategy has been implemented in programming sketching education, where users are required to pause after each attempt and complete a questionnaire comprising reflective questions. This structured pause facilitates the transition from active problem-solving to reflection phases [9].

2) Changing Habitual Thinking Patterns: Breaking established thinking habits often requires the introduction of alternative approaches [31]. For instance, Shrestha et al. [20] demonstrated this by deliberately limiting robot capabilities to prevent the processing of ambiguous command dialogues, thereby training users to develop more precise sentence construction skills.

3) Learning from Falsifications and Feedback: Metacognitive development is significantly enhanced through error recognition and feedback processing. Shrestha et al. [20] implemented a system requiring users to engage in iterative trial-and-error processes to decompose actions and provide executable instructions to robots, learning from previous errors to develop programming thinking. Similarly, Kong and Yang [32] developed an ”Error-feedback” system where students formulated and tested hypotheses while observing robot errors during training, such as wall collisions and navigation issues,

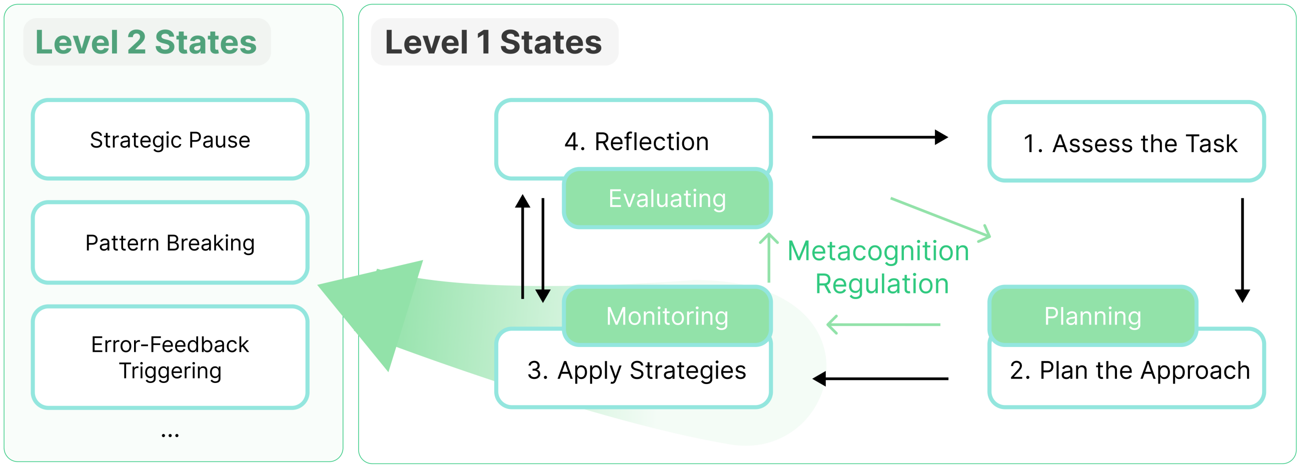

Fig. 1. Two-level state-based framework for robot interfaces. Level 1 states show the metacognition regulation cycle, while Level 2 states provide specific intervention strategies that activate during strategy application.

enabling them to adjust strategies and make corrections based on received feedback.

III. Design

A. Design Implications

Based on our analysis of metacognitive processes in human-robot interaction, we propose several key design implications for developing effective interfaces that support metacognitive regulation. These implications focus on three primary aspects of interface functionality:

- First, the interface should effectively convey explicit metacognitive information that is mutually understandable between human senders and robot receivers.
- Second, the interface should provide instructions and strategic guidance during metacognitive monitoring through interventional strategies to facilitate users' metacognitive development.
- Third, the interface should adapt dynamically to both human and robot actions during the metacognition formulation process.

B. Proposed Approach

The human-robot interaction framework is designed to facilitate metacognitive development during task execution. The interface design incorporates metacognitive regulation theory (see II-B) by explicitly modeling and representing the user's metacognitive processes. This mapping enables the system to support and enhance human metacognition through interactive feedback mechanisms and process visualization. We propose a two-level state-based framework for robot interface design that supports metacognitive regulation (see Figure 1). The framework consists of Level 1 states representing the metacognition circle during task execution, and Level 2 states indicating specific metacognitive strategies. Figure 1 illustrates this hierarchical structure, where Level 1 states form a continuous cycle of metacognitive regulation, while Level 2 states provide strategic interventions during the process. The Level 1 states correspond to the phases of metacognitive regulation during task execution:

1) State 1: Assess the Task: In this initial state, the robot interface serves as a medium for task demonstration. The interface presents tasks to users through various forms such as games or quizzes, focusing purely on information presentation without intervention.

2) State 2: Plan the Approach: During the planning phase, the interface provides instructional guidance about potential solution approaches, including specific knowledge requirements and applicable methods. This state involves task-specific knowledge that varies according to different contexts and requirements.

3) State 3: Apply Strategies: This state implements metacognitive strategies selected from Level 2 to facilitate users' metacognitive processes. The interface actively engages users through various strategic interventions, including strategic pauses, changing habitual thinking patterns, and learning from falsifications and feedback.

4) State 4: Reflection: The reflection state features interactive functionalities where users can record and express their understanding of the task. This is the primary state for receiving user input, and the interface updates based on user-provided information and knowledge from previous reflection phases. This iterative process ensures continuous metacognitive development throughout the task execution cycle. State 3 and State 4 can transition between each other.

The Level 2 states directly correspond to the metacognitive monitoring strategies discussed in Section II-C: Strategic Pause, Pattern Breaking, and Error-Feedback Triggering. These states provide specific interventional strategies to implement during the Level 1 cycle, particularly during the Apply Strategies phase(III-B3).

C. A User example

To illustrate our design framework, we present an example of a second language learning task, following the Level 1 cycle. The robot interface begins by presenting a task focused on acquiring fruit-related vocabulary. It then provides instructional guidance on potential strategies for task completion, such as listening to pronunciations, repeating after the robot, and engaging in matching exercises. As the learner progresses, the robot actively supports their efforts through strategic interventions, transitioning to Level 2 states when necessary. For example, it corrects the learner's mispronunciations, facilitating the change of habitual thinking patterns as described in Section II-C. Finally, the interface enables learners to record and reflect on their progress, allowing them to adapt their strategies to better address their needs in future sessions.

IV. Limitations and Future Work

Our current research has several limitations that suggest directions for future investigation. In-depth user studies are necessary to evaluate the flexibility and effectiveness of metacognition theory-guided human-robot interactions. Our method requires further comprehensive research to validate the adaptability of the HRI framework across different task domains [5]. Particular attention should be paid to how humans acquire task-specific metacognitive knowledge in various contexts. These studies would help validate our theoretical framework and provide practical insights for implementation. Moreover, technical implementation details require further exploration. This includes investigating how different interface states can be implemented, determining conditions for state transitions, and developing algorithms that enable robots to process information and generate appropriate interfaces more effectively. While our study primarily focused on metacognition regulation during human-robot interaction, future research should examine metacognition in long-term behavioral change. Specifically, investigation is needed into how humans maintain surveillance over their beliefs to ensure consistent alignment with goals. This includes incorporating aspects of motivation, belief systems, and confidence [33], which are also important factors affecting human metacognition.

References

[1] K. J. Rohlfing, N. Altvater-Mackensen, N. Caruana, R. Van den Berghe, B. Bruno, N. F. Tolksdorf, and A. Hanulíková, "Social/dialogical roles of social robots in supporting children's learning of language and literacy," *Frontiers in Robotics and AI*, vol. 9, p. 971749, 2022.

[2] I. Brinck and R. Liljenfors, "The developmental origin of metacognition," *Infant and Child Development*, vol. 22, no. 1, pp. 85–101, 2013.

[3] A. Brown, "Metacognition, executive control, self-regulation, and other more mysterious mechanisms," *Metacognition, Motivation and Understanding*, pp. 65–116, 1987.

[4] H. Kallio, K. Virta, and M. Kallio, "Modelling the components of metacognitive awareness," *International Journal of Educational Psychology*, vol. 7, no. 2, pp. 94–122, 2018.

[5] S. B. Özkan and M. Toz, "A review study on the investigation of the effects of using robots in education on metacognitive behaviors," *Computer Applications in Engineering Education*, vol. 30, no. 4, pp. 1277–1288, 2022.

[6] T. J. Prescott, K. Vogeley, and A. Wykowska, "Understanding the sense of self through robotics," *Science Robotics*, vol. 9, no. 95, p. eadn2733, 2024.

[7] D. Pnevmatikos, P. Christodoulou, and N. Fachantidis, "Designing a socially assistive robot for education through a participatory design approach," *International Journal of Social Robotics*, vol. 14, no. 3, pp. 763–788, 2022.

[8] S. De Felice, A. F. d. C. Hamilton, M. Ponari, and G. Vigliocco, "Learning from others is good, with others is better: The role of social interaction in human acquisition of new knowledge," *Philosophical Transactions of the Royal Society B*, vol. 378, no. 1870, p. 20210357, 2023.

[9] M. Havenga and T. Jordaan, "Introducing coding and robotics: Prospective mathematics teachers' metacognitive thinking in scratch," in *International Conference on Robotics in Education*. Springer, 2023, pp. 17–26.

[10] A. Kouzalis, A. Antoniou, N. Rossides, R. Panaoura, and P. Yadav, "Advanced technologies and mathematical metacognition: The present and future orientation," *BioSystems*, vol. 245, p. 105312, 2024.

[11] C. Socratous and A. Ioannou, "Evaluating the impact of the curriculum structure on group metacognition during collaborative problem-solving using educational robotics," *TechTrends*, vol. 66, no. 5, pp. 771–783, 2022.

[12] D. Mishra, R. G. Lugo, K. Parish, and S. Tilden, "Metacognitive processes involved in human robot interaction in the school learning environment," in *International Conference on Human-Computer Interaction*. Springer, 2023, pp. 85–100.

[13] R. M. Williams, K. Alikhademi, I. N. S. Munyaka, and J. E. Gilbert, "Metacogs: Mitigating executive dysfunction via agent-based modeling for metacognitive strategy development," *ACM Transactions on Accessible Computing*, vol. 15, no. 3, pp. 1–32, 2022.

[14] M. Romero and G. Kalmpourtzis, "Metacognition and self-regulated learning in a manipulative robotic problem-solving task," in *Metacognition and Education: Future Trends*. Routledge, 2023, pp. 105–123.

[15] M. M. Eshtehard, N. Babakhani, and P. Hassani Abharian, "The effect of training based on metacognition and emotion regulation on health behavior self-management of overweight middle-aged people," *Practice in Clinical Psychology*, vol. 11, no. 3, pp. 201–210, 2023.

[16] C. Socratous, "Collaborative knowledge construction and metacognition via educational robotics," Ph.D. dissertation, Cyprus University of Technology, 2022.

[17] L. Peternel, N. Tsagarakis, and A. Ajoudani, "Towards multi-modal intention interfaces for human-robot co-manipulation," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2016, pp. 2663–2669.

[18] L. Yuan, X. Gao, Z. Zheng, M. Edmonds, Y. N. Wu, F. Rossano, H. Lu, Y. Zhu, and S.-C. Zhu, "In situ bidirectional human-robot value alignment," *Science Robotics*, vol. 7, no. 68, p. eabm4183, 2022.

[19] J. Del Olmo-Muñoz, P. D. Diago, D. Arnau, D. Arnau-Blasco, and J. A. González-Calero, "Metacognitive control during problem solving at early ages in programming tasks using a floor robot," *ZDM–Mathematics Education*, vol. 56, no. 6, pp. 1303–1317, 2024.

[20] A. Shrestha, R. Ishikawa, M. Sano, and S. Ooi, "Research on support system for programmatic thinking based on metacognition of instructional interaction," in *International Conference on Information and Education Technology (ICIET)*. IEEE, 2024, pp. 253–257.

[21] G. Schraw, K. J. Crippen, and K. Hartley, "Promoting self-regulation in science education: Metacognition as part of a broader perspective on learning," *Research in Science Education*, vol. 36, pp. 111–139, 2006.

[22] C. D. Frith, "The role of metacognition in human social interactions," *Philosophical Transactions of the Royal Society B*, vol. 367, no. 1599, pp. 2213–2223, 2012.

[23] C. Heyes, D. Bang, N. Shea, C. D. Frith, and S. M. Fleming, "Knowing ourselves together: The cultural origins of metacognition," *Trends in Cognitive Sciences*, vol. 24, no. 5, pp. 349–362, 2020.

[24] N. Shea, A. Boldt, D. Bang, N. Yeung, C. Heyes, and C. D. Frith, "Supra-personal cognitive control and metacognition," *Trends in Cognitive Sciences*, vol. 18, no. 4, pp. 186–193, 2014.

[25] J. Dunstone and C. A. Caldwell, "Cumulative culture and explicit metacognition: A review of theories, evidence and key predictions," *Palgrave Communications*, vol. 4, no. 1, 2018.

[26] M. F. Teng, C. Qin, and C. Wang, "Validation of metacognitive academic writing strategies and the predictive effects on academic writing performance," *Metacognition and Learning*, vol. 17, no. 1, pp. 167–190, 2022.

[27] A. Desoete, H. Roeyers, and A. Buysse, "Metacognition and mathematical problem solving in grade 3," *Journal of Learning Disabilities*, vol. 34, no. 5, pp. 435–447, 2001.

[28] K. Pisessith, S. Lertyongphati, and R. J. Rogers, *The Quick Guide to Classroom Management: 45 Secrets That All High School Teachers Need to Know*. Amazon Digital Services, 2021.

[29] M. Pressley, J. G. Borkowski, and W. Schneider, "Cognitive strategies: Good strategy users coordinate metacognition and knowledge," *Annals of Child Development*, vol. 4, pp. 89–129, 1987.

[30] V. Bamicha and A. Drigas, "Human-social robot interaction in the light of ToM and metacognitive functions," *Scientific Electronic Archives*, vol. 17, no. 5, 2024.

[31] D. Disalvo, *Brain Changer: How Harnessing Your Brain's Power to Adapt Can Change Your Life*. BenBella Books, 2013.

[32] S.-C. Kong and Y. Yang, "Using the robot-assisted attention-engagement-error-feedback-reflection (AEER) pedagogical design to develop machine learning concepts," in *International Conference on Computer Supported Education*. SCITEPRESS, 2024, pp. 155–162.

[33] A. Mazancieux, M. Pereira, N. Faivre, P. Mamassian, C. J. A. Moulin, and C. Souchay, "Towards a common conceptual space for metacognition in perception and memory," *Nature Reviews Psychology*, vol. 2, no. 12, pp. 751–766, 2023.