Towards Supporting Mediators in Human–Agent Collaboration

Anirban Mukhopadhyay
Virginia Tech
Department of Computer Science
Virginia, USA

Kurt Luther
Virginia Tech
Department of Computer Science
Virginia, USA

Abstract

AI agents are increasingly positioned as collaborators in team settings, but their success depends on more than task performance. In this position paper, we argue that effective human–agent collaboration requires supporting the underlying social and cognitive processes that enable teamwork. Building on our empirical study of proactive AI agents in small-team, time-sensitive collaborative tasks, we highlight recurring challenges related to intervention content, timing, and position. We frame these challenges through the Input-Mediator-Output-Input (IMOI) model from organizational psychology, which emphasizes mediating processes such as trust, safety, shared mental models, transactive memory, and communication among others. We show how breakdowns in these mediators explain why agents sometimes disrupt collaboration despite strong technical capabilities. We outline design considerations that can position agents as successful remote collaborators by supporting team processes like planning, structuring, and adaptation.

ACM Reference Format:
Anirban Mukhopadhyay and Kurt Luther. 2026. Towards Supporting Mediators in Human–Agent Collaboration. In *Proceedings of (CHI '26).* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 Introduction

Human–agent collaboration has been studied across domains such as decision making, creative work, and problem solving [5, 11, 15]. AI agents are expected to function as teammates that contribute to shared goals rather than simply respond to individual prompts [1, 13, 14]. However, enabling agents to collaborate effectively with humans remains challenging. Agents that are technically capable may still fail as collaborators when they disrupt coordination, trust, or shared understanding within a group.

We argue that successful human–agent collaboration requires moving beyond performance-oriented perspectives toward a process-centered view of teamwork. Organizational and CSCW research has long shown that team success depends on mediating processes such as planning, communication, adaptation, and workload distribution [6, 8]. These processes become even more fragile in hybrid human–agent settings, where agents may intervene without awareness of conversational flow or team dynamics [3, 7, 10]. Supporting these team processes is particularly critical if AI agents are to operate as effective remote collaborators in real-world teamwork scenarios.

To ground this argument empirically, we draw on insights from our recent CHI 2026 study [12]. In that work, we investigated how proactive generative AI agents influence small-team collaboration under time pressure. The study setup is presented in Figure 1. In a within-subjects mixed-methods study, four-person teams solved escape-room–style puzzles while interacting with two distinct AI roles embedded directly into co-located teamwork. A Facilitator agent scaffolded coordination through summaries and group structure prompts, while a Peer agent proactively contributed ideas and responded to follow-up queries. Unlike reactive tools, both agents intervened autonomously during live collaboration. Beyond performance outcomes, we examined how these interventions shaped coordination, reliance, workload, and communication patterns.

The results revealed a wide variation in agent perceptions and puzzle solving performance. The peer agent occasionally improved problem-solving by offering timely hints and supporting memory offloading. However, it also increased subjective workload, disrupted conversational flow, and sometimes led to over-reliance and siloed interaction. Teams followed divergent trajectories where initial curiosity sometimes shifted toward dependence, frustration, and disengagement. The facilitator agent helped structure early collaboration and anchor shared focus, but its summaries were often perceived as redundant or poorly timed, leading to marginalization over time. These findings highlight that effectiveness of human–agent collaboration depends not only on what the agent contributes, but also on when it intervenes and the role it plays within the team.

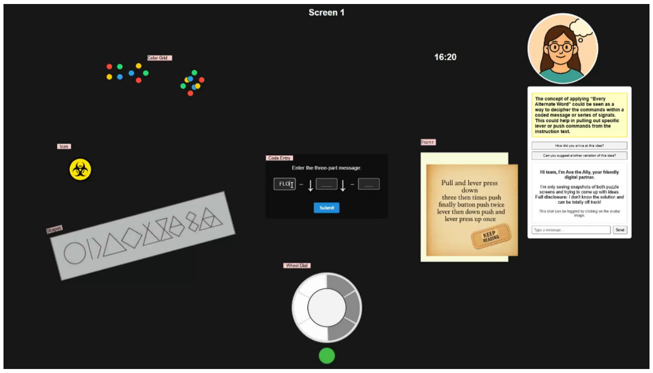
2 A Process-Centered Lens To Frame The Challenges

To understand these breakdowns, we situate human–agent collaboration within established CSCW and organizational theories of teamwork in this position paper. The IMOI (Input-Mediator-Output-Input) framework provides a lens for analyzing team effectiveness [6, 9]. Rather than focusing only on inputs (e.g., system capabilities, team composition) or outputs (e.g., performance), IMOI emphasizes mediators—the cognitive, social, and behavioral processes that explain why teams succeed or fail.

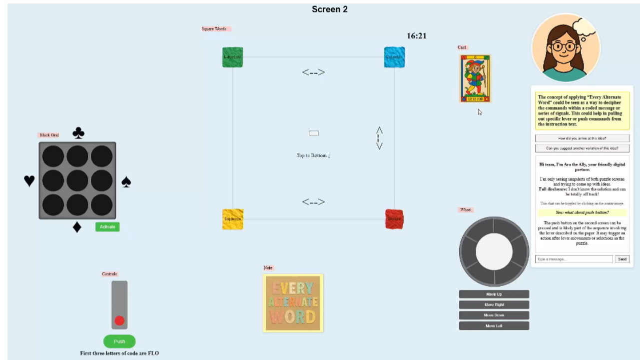
The framework identifies three phases: (1) Forming (IM), involving trust, planning, and structuring; (2) Functioning (MO), involving adaptation, bonding, and learning; and (3) Finishing (OI), involving knowledge transfer and reflection. Recent work has extended IMOI to human–AI teams, showing that mediating processes such as shared mental models, transactive memory systems, and psychological safety are often more fragile in hybrid teams than in human-only settings [2, 4, 9]. We posit that for an AI agent to be

CHI '26, Barcelona, Spain

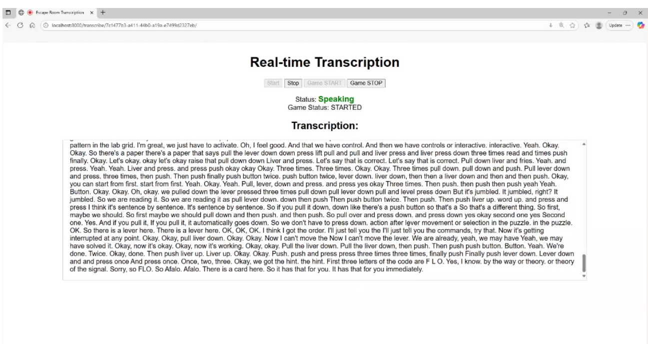
ACM ISBN 978-x-xxxx-xxxx-x/YYYY/MM
https://doi.org/10.1145/nnnnnnn.nnnnnnn



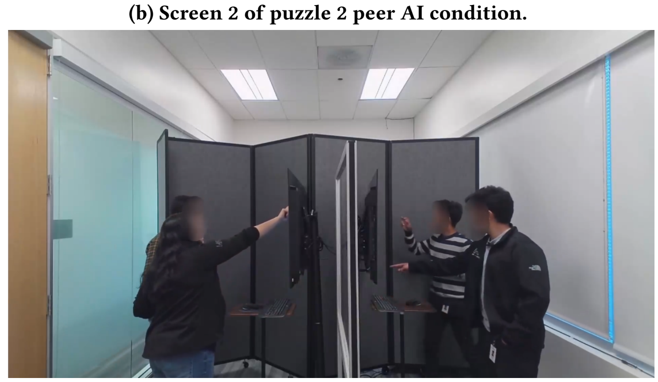
(a) Screen 1 of puzzle 2 peer AI condition.



(b) Screen 2 of puzzle 2 peer AI condition.



(c) Real-time transcription of group dialogue.



(d) Group interaction during puzzle 2 task

Figure 1: Recording of Group 6, Session 2 (Puzzle 2 with the peer agent). The top panels, (a) and (b), capture the two puzzle screens, each displayed on separate TVs connected to laptops. (c) Each participant wore a lavalier microphone, and their audio was merged into a single channel and transcribed in real time using WhisperX. These transcripts were used by the peer agent to contextualize responses and by the facilitator to generate summaries. (d) Finally, a room camera captured group interactions and overall activity throughout the session.

a successful remote collaborator, it must support these mediating team processes.

In this workshop paper, we focus specifically on a few aspects of the Forming and Functioning phases as these stages capture the most immediate breakdowns observed in our empirical work. While the Finishing phase (e.g., long-term learning and transfer) is equally important, stabilizing early coordination and real-time adaptation is a necessary first step for agents operating as collaborators.

3 Agent Features for Supporting Team Processes

3.1 Supporting Planning and Information Coordination (Forming)

Planning in human–agent teams is fragile, particularly under time pressure. Teams must identify relevant information, coordinate contributions, and develop shared strategy. In our study, teams frequently lost track of which ideas were generated by whom, what information had been surfaced, and which hypotheses were still unresolved. Agent interventions sometimes exacerbated this fragmentation when contributions disregarded conversational context.

To function as effective collaborators, agents must be able to model and align with team planning processes. This includes maintaining awareness of division of work, tracking unresolved goals, and selectively contributing when informational gaps are detected. Rather than injecting generic suggestions, agents should support collective planning by reinforcing shared task representations and clarifying open questions.

3.2 Supporting Structuring: Shared Mental Models and Transactive Memory (Forming)

Effective structuring depends on two foundational processes: (1) shared mental models (SMM), a common understanding of task goals and strategies; and (2) transactive memory systems (TMS), knowing who knows what within the team. In human–agent teams, these processes extend to the agent itself. Team members must understand the agent's capabilities and limitations, and the agent must maintain an evolving model of team expertise and roles.

Our findings show that miscalibration occurred when teams lacked clarity about the agent's competence or role. Sometimes the agent was treated as an authoritative expert; at other times it was ignored. These oscillations indicate unstable shared mental models.

To address this, agents must explicitly communicate their scope, uncertainty, and reasoning context in ways that support trust calibration. Additionally, agents should maintain and surface models of team expertise, enabling better task delegation and reducing redundant effort.

3.3 Supporting Adaptation and Workload Sharing (Functioning)

During execution, teams must adapt to changing conditions and redistribute effort. In our study, agents sometimes added cognitive load, particularly when interventions interrupted human coordination. Teams also struggled to make sense of agent output and formulate effective prompts.

Agents intended for collaborative settings must therefore support adaptive coordination. This includes sensitivity to timing, awareness of conversational flow, and the ability to modulate intervention frequency. Agents should dynamically adjust their level of proactivity based on team state—for example, offering support during stalls but deferring when human coordination is fluid. Effective workload sharing also requires that agents signal their strengths and limitations clearly.

4 Implications for Human–Agent Collaboration Research

From a CSCW perspective, these ideas reinforce the importance of treating AI agents as participants in collaborative systems rather than tools in workflows. The success of a remote AI collaborator depends on its ability to sustain mediating team processes like planning, structuring, and adaptation.

We argue that future human–agent collaboration research should move toward process-aware agent design, where agents are evaluated not only by task accuracy but by how they shape shared understanding, coordination patterns, and team dynamics over time. Supporting these mediators is especially critical in remote and distributed contexts, where breakdowns in shared awareness are amplified.

References

[1] Chun-Wei Chiang, Zhuoran Lu, Zhuoyan Li, and Ming Yin. 2023. Are Two Heads Better Than One in AI-Assisted Decision Making? Comparing the Behavior and Performance of Groups and Individuals in Human-AI Collaborative Recidivism Risk Assessment. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, 1–18. doi:10.1145/3544548.3581015

[2] Christopher Flathmann, Nathan J McNeese, and Thomas A O'Neill. 2024. Designing High-Impact Experiments for Human–Autonomy/AI Teaming. *Journal of Cognitive Engineering and Decision Making* (2024), 15553434251327697.

[3] Tessa Haesevoets, David De Cremer, Kim Dierckx, and Alain Van Hiel. 2021. Human-machine collaboration in managerial decision making. *Computers in Human Behavior* 119 (June 2021), 106730. doi:10.1016/j.chb.2021.106730

[4] Nader Hanna and Deborah Richards. 2018. The Impact of Multimodal Communication on a Shared Mental Model, Trust, and Commitment in Human–Intelligent Virtual Agent Teams. *Multimodal Technologies and Interaction* 2, 3 (Aug. 2018). doi:10.3390/mti2030048 Company: Multidisciplinary Digital Publishing Institute Distributor: Multidisciplinary Digital Publishing Institute Institution: Multidisciplinary Digital Publishing Institute Label: Multidisciplinary Digital Publishing Institute Publisher: publisher.

[5] Angel Hsing-Chi Hwang, John Oliver Siy, Renee Shelby, and Alison Lentz. 2024. In whose voice?: examining AI agent representation of people in social interaction through generative speech. In *Proceedings of the 2024 ACM Designing Interactive Systems Conference*. 224–245.

[6] Daniel R. Ilgen, John R. Hollenbeck, Michael Johnson, and Dustin Jundt. 2005. Teams in Organizations: From Input-Process-Output Models to IMOI Models. *Annual Review of Psychology* 56, Volume 56, 2005 (Feb. 2005), 517–543. doi:10.1146/annurev.psych.56.091103.070250 Publisher: Annual Reviews.

[7] N. R. Jennings, L. Moreau, D. Nicholson, S. Ramchurn, S. Roberts, T. Rodden, and A. Rogers. 2014. Human-agent collectives. *Commun. ACM* 57, 12 (Nov. 2014), 80–88. doi:10.1145/2629559

[8] Norbert L. Kerr and R. Scott Tindale. 2004. Group Performance and Decision Making. *Annual Review of Psychology* 55, Volume 55, 2004 (Jan. 2004), 623–655. doi:10.1146/annurev.psych.55.090902.142009 Publisher: Annual Reviews.

[9] Styliani Kleanthous. 2024. Human-AI Teaming: Following the IMOI Framework. In *Artificial Intelligence in HCI*, Helmut Degen and Stavroula Ntoa (Eds.). Springer Nature Switzerland, Cham, 387–406. doi:10.1007/978-3-031-60611-3_27

[10] Vivian Lai, Chacha Chen, Alison Smith-Renner, Q. Vera Liao, and Chenhao Tan. 2023. Towards a Science of Human-AI Decision Making: An Overview of Design Space in Empirical Human-Subject Studies. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency* (Chicago, IL, USA) *(FAccT '23)*. Association for Computing Machinery, New York, NY, USA, 1369–1385. doi:10.1145/3593013.3594087

[11] Meredith Ringel Morris, Carrie J. Cai, Jess Holbrook, Chinmay Kulkarni, and Michael Terry. 2023. The Design Space of Generative Models. doi:10.48550/arXiv.2304.10547 arXiv:2304.10547 [cs].

[12] Anirban Mukhopadhyay, Kevin Salubre, Hifza Javed, Shashank Mehrotra, and Kumar Akash. 2026. Exploring The Impact Of Proactive Generative AI Agent Roles In Time-Sensitive Collaborative Problem-Solving Tasks. (2026). https://anmukhop.github.io/uploads/CHI2026-Paper.pdf

[13] Simone van den Broek, Supraja Sankaran, Jan de Wit, and Alwin de Rooij. 2024. Exploring the Supportive Role of Artificial Intelligence in Participatory Design: A Systematic Review. In *Proceedings of the Participatory Design Conference 2024: Exploratory Papers and Workshops - Volume 2 (PDC '24, Vol. 2)*. Association for Computing Machinery, New York, NY, USA, 37–44. doi:10.1145/3661455.3669868

[14] Mengyao Wang, Jiayun Wu, Shuai Ma, Nuo Li, Peng Zhang, Ning Gu, and Tun Lu. 2025. Adaptive Human-Agent Teaming: A Review of Empirical Studies from the Process Dynamics Perspective. doi:10.48550/arXiv.2504.10918 arXiv:2504.10918 [cs].

[15] Rui Zhang, Wen Duan, Christopher Flathmann, Nathan McNeese, Guo Freeman, and Alyssa Williams. 2023. Investigating AI Teammate Communication Strategies and Their Impact in Human-AI Teams for Effective Teamwork. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW2 (Sept. 2023), 1–31. doi:10.1145/3610072