

AI and Group Decision Making: An Information Processing Perspective

Thomas E. Gorman Torsten Reimer

Introduction

- (Steyvers & Kumar, 2024)
- (Lai et al., 2023)
- (Burton et al., 2024)
- (Rastogi et al., 2023)

....

Inputs

..

Group Member Roles

Deciding how best to assign team members to roles is crucial in group decision-making, particularly when learning who is best suited for what role within a team. Marjeh et al. (2024) explore how humans allocate tasks within teams comprising both human and AI agents to maximize overall performance. The central theme of their research is understanding the mechanisms by which individuals discern and act upon their own strengths and those of their team members in a dynamic task allocation setting. In their experimental paradigm, participants had to repeatedly allocate three different types of tasks (visual,

auditory, and lexical tasks) between themselves and two AI agents. Unbeknownst to participants, each AI agent was configured to have high competence (70% success rate) in one task type but low competence (15% success rate) in others.

Recent advances in large language models have dramatically expanded the potential roles of AI in group decision-making, enabling AI agents to move beyond simple advisory functions to serve as mediators, devil’s advocates, and active discussion participants

Chiang et al. (2024) investigated the potential of Large Language Models (LLMs) to act as devil’s advocates in AI-assisted group decision-making - in the hopes of fostering more critical engagement with AI assistance. In their experimental task, participants were first individually trained on the relationship between defendant profiles and recidivism. For each defendant, participants were also shown the prediction of a recommendation AI model (RiskComp). Participants were then sorted into groups of three, where they reviewed and discussed novel defendant profiles, before making a group recidivism assessment. In the group stage, the recommendations from the RiskComp model were biased against a subset of the defendants (black defendants with low prior crime counts). Of interest was whether the inclusion of an LLM-based devil’s advocate in the group discussions could help mitigate the bias introduced by the RiskComp AI model (note that the LLM devils advocate and RiskComp AI are separate AI models). The experimental manipulation consisted of four variants of an LLM-based devil’s advocate using, varying both the target of objection (challenging either RiskComp recommendations or majority group opinions) and the level of interactivity (static one-time comments versus dynamic engagement throughout the discussions). Their findings revealed that the dynamic devil’s advocate led to higher decision accuracy and improved discernment of when to trust the RiskComp model’s advice.

- (Marjeh et al., 2024)
- (Kumar et al., 2024)

- (Lu et al., 2024)
- (McNeese et al., 2023)

Information Processing

Information Search

- (Gao & Zhang, 2024)

...

Communication; information sharing

Transactive memory systems (TMS) represent a critical aspect of group cognition, referring to the shared understanding within a group regarding the distribution of knowledge and expertise among its members (Wegner, 1987). A well-functioning TMS enables team members not only to know who possesses specific knowledge but also to access and share this distributed expertise efficiently.

Bienefeld et al. (2023) conducted an observational study to examine the role of transactive memory systems and speaking-up behaviors in human-AI teams within an intensive care unit (ICU) setting. In this study, ICU physicians and nurses, divided into groups of four, who collaborated with an AI agent named “Autovent.” Autovent is an auto-adaptive ventilator system that autonomously manages patient ventilation by processing continuous, individualized data streams. Participants, all with a minimum of six months’ experience using Autovent, engaged in simulated clinical scenarios that required diagnosing and treating critically ill patients. Using behavioral coding of video recordings, the researchers analyzed how team members accessed information from both human teammates and the AI system, investigating how these human-human and human-ai interactions related to subsequent behaviors like hypothesis generation and speaking up with concerns. The researchers found that in higher-performing teams, accessing knowledge from the AI agent was positively correlated with developing new hypotheses and increased speaking-up behavior. Conversely,

accessing information from human team members was negatively associated with these behaviors, regardless of team performance. These results suggest that AI systems may serve as unique knowledge repositories that help teams overcome some of the social barriers that typically inhibit information sharing and voice behaviors in purely human teams.

- (Yang et al., 2024)
- (Ma et al., 2024)
- (Radivojevic et al., 2024)
- (Sidji et al., 2024)
- (Nishida et al., 2024)
- (Chuang et al., 2024)

Shared Mental Models

- (Collins et al., 2024)

...

Cognitive Load

Buçinca et al. (2021) examined how interface design might influence cognitive engagement with AI recommendations through what they term “cognitive forcing functions.” Drawing on dual-process theory, they implemented three distinct interface interventions (e.g., requiring explicit requests for AI input, mandating initial independent decisions, introducing temporal delays) designed to disrupt automatic processing and promote more analytical engagement with AI suggestions. Their findings demonstrated that while these interventions successfully reduced overreliance on incorrect AI recommendations, they also increased perceived cognitive load and decreased user satisfaction. Of particular methodological interest was their systematic investigation of individual differences in cognitive motivation: participants with high Need for Cognition (NFC) showed substantially greater benefits from these interventions, suggesting that the effectiveness of such cognitive load manipulations may be moderated by individual differences in information processing preferences.

Decision-Making Output

..

Consensus Formation

Tessler et al. (2024) investigated the potential of AI in facilitating consensus formation through their development of the “Habermas Machine” (HM), an LLM-based system fine-tuned to mediate human deliberation. The HM system receives input statements from individual participants, and attempts to generate consensus statements which will maximize group endorsement. The findings revealed that the AI-generated group statements were consistently preferred over comparison statements written by human mediators. Participants rated the AI-mediated statements higher in terms of informativeness, clarity, and lack of bias. This suggests that AI can effectively capture the collective sentiment of a group and articulate it in a way that resonates with its members. Notably, the researchers also verified that the HM system reliably incorporated minority opinions into the consensus statements, preventing dominance by majority perspectives. These results were replicated in a virtual citizens’ assembly with a demographically representative sample of the UK population. The AI-mediated process again resulted in high-quality group statements and facilitated consensus among participants on contentious issues.

Decision Accuracy and Confidence

- (Becker et al., 2022)
- ...

Trust, Risk and Reliance

Trust in AI

- Westphal et al. (2023)
- (Koehl & Vangsness, 2023)
- (Banerjee et al., 2024)

Reliance

- (Narayanan et al., 2023)

Recent work has begun examining how people attribute responsibility in human-AI collaborative contexts where control is shared and actions are interdependent (Tsirtsis et al., 2024). Their study employs a stylized semi-autonomous driving simulation where participants observe how a ‘human agent’ and an ‘AI agent’ collaborate to reach a destination within a time limit. In their setup, the human and AI agents shared control of a vehicle, with each agent having partial and differing knowledge of the environment (i.e., the AI knew about traffic conditions but not road closures, while humans knew about closures but not traffic). Participants observe illustrated simulations of a variety of commute scenarios, and then make judgements about how responsible each agent was for the commute outcome (reaching the destination on time, or not). The study reveals that participants’ responsibility judgments are influenced by factors such as the unexpectedness of an agent’s action, counterfactual simulations of alternative actions, and the actual contribution of each agent to the task outcome.

Utilization

Recent work by Bućinca et al. (2021) presents an innovative approach to addressing overreliance on AI systems through interface design rather than explanation quality. Their study evaluated three “cognitive forcing functions” - interface elements designed to disrupt quick, heuristic processing of AI recommendations. Although these interventions significantly reduced overreliance on incorrect AI recommendations, an important trade-off emerged: interfaces that most effectively prevented overreliance were also rated as most complex and least preferred by users. Moreover, their analysis revealed potential equity concerns, as the interventions provided substantially greater benefits to individuals with high Need for Cognition. These findings suggest that while interface design can effectively modulate AI utilization patterns, careful consideration must be given to both

user experience and potential intervention-generated inequalities.

- (Cui & Yasseri, 2024)
- (Stadler et al., 2024)

Risk

- (Bhatia, 2024)
- (Zhu et al., 2024)

Tables

Table 1: Research Methods in Human-AI Group Decision-Making

Research Method	Task Types	Key Measures	Example Studies	Strengths	Limitations
Laboratory Experiments	Simplified, controlled tasks (e.g., classification, judgment, resource allocation, logic puzzles)	Decision accuracy, response time, trust ratings, reliance on AI, eye-tracking, think-aloud protocols, physiological measures	Chiang et al. (2024); Bućinca et al. (2021); Swaroop et al. (2024)	High control, causal inference, precise measurement of behavior	Limited ecological validity, often simplified tasks, potential for demand characteristics, limited generalizability
Field Studies	Complex, real-world tasks (e.g., medical diagnosis, emergency response, project planning, incident response)	Team performance, communication patterns, system adoption, qualitative observations (interviews, focus groups), workflow analysis	Bienefeld et al. (2023); Yang et al. (2024)	High ecological validity, real-world outcomes	Less experimental control, smaller sample sizes, difficult to isolate specific AI effects, potential for confounding variables
Online Experiments	Varied tasks (e.g., text analysis, opinion formation, consensus building, games, surveys, simulated work tasks)	Group agreement, information sharing, bias measures, survey responses, behavioral logs, communication content analysis	Tessler et al. (2024); Chuang et al. (2024); Sidji et al. (2024); Nishida et al. (2024)	Large samples, cost-effective, access to diverse populations, can study large-scale dynamics	Less control over participation, limited interaction depth, potential for self-selection bias, difficulty ensuring data quality
Observational Studies	Real-world collaborative activities (e.g., collaborative writing, decision support usage, team coordination, online discussions)	Usage patterns, adaptation over time, natural behaviors, communication analysis, qualitative observations	Radivojevic et al. (2024); Ma et al. (2024)	Natural behavior, temporal dynamics, rich qualitative data, can study evolving interactions	No experimental manipulation, selection effects, difficulty in establishing causality, observer bias

Table 2: Types of AI Used in Decision-Making Tasks

AI Type	Key Capabilities	Limitations	Primary Applications	Representative Studies
Large Language Models	Natural language understanding; Open-ended dialogue; Context adaptation; Knowledge synthesis; Abstract reasoning	Hallucinations; Inconsistency; Limited causal reasoning; Training data cutoff	Group facilitation; Information synthesis; Creative ideation; Deliberation support	Tessler et al. (2024); Chiang et al. (2024); Collins et al. (2024); Du et al. (2024)
Domain-Specific AI	Expert knowledge; Precise calculations; Consistent performance; Real-time processing; Specialized optimization	Limited generalization; Rigid decision rules; Poor adaptation to novel scenarios	Medical diagnosis; Resource allocation; Risk assessment; Pattern detection	Bienefeld et al. (2023); Marjeh et al. (2024); Yang et al. (2024)
Hybrid Systems	Combined generalist/specialist capabilities; Integrated expertise; Adaptive reasoning; Multi-modal processing	Integration complexity; System conflicts; Higher computational needs	Complex decision-making; Strategic planning; Knowledge management	Burton et al. (2024); Ma et al. (2024); Gao et al. (2024)
Autonomous Agents	Independent operation; Goal-directed behavior; Environmental interaction; Adaptive learning	Limited social awareness; Trust-building challenges; Coordination issues	Team collaboration; Process automation; Distributed tasks	McNeese et al. (2023); Guo et al. (2024); Zhang et al. (2024)

References

- Banerjee, D., Teso, S., Sayin, B., & Passerini, A. (2024). *Learning To Guide Human Decision Makers With Vision-Language Models* (arXiv:2403.16501). arXiv. <https://arxiv.org/abs/2403.16501>
- Becker, F., Skirzyński, J., van Opheusden, B., & Lieder, F. (2022). Boosting Human Decision-making with AI-Generated Decision Aids. *Computational Brain & Behavior*, 5(4), 467–490. <https://doi.org/10.1007/s42113-022-00149-y>
- Bhatia, S. (2024). Exploring variability in risk taking with large language models. *Journal of Experimental Psychology: General*, 153(7), 1838–1860. <https://doi.org/10.1037/xge0001607>
- Bienefeld, N., Kolbe, M., Camen, G., Huser, D., & Buehler, P. K. (2023). Human-AI teaming: Leveraging transactive memory and speaking up for enhanced team effectiveness. *Frontiers in Psychology*, 14. <https://doi.org/10.3389/fpsyg.2023.1208019>
- Buçinca, Z., Malaya, M. B., & Gajos, K. Z. (2021). To Trust or to Think: Cognitive Forcing Functions Can Reduce Overreliance on AI in AI-assisted Decision-making. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW1), 1–21. <https://doi.org/10.1145/3449287>
- Burton, J. W., Lopez-Lopez, E., Hechtlinger, S., Rahwan, Z., Aeschbach, S., Bakker, M. A., Becker, J. A., Berditchevskaia, A., Berger, J., Brinkmann, L., Flek, L., Herzog, S. M., Huang, S., Kapoor, S., Narayanan, A., Nussberger, A.-M., Yasseri, T., Nickl, P., Almaatouq, A., ... Hertwig, R. (2024). How large language models can reshape collective intelligence. *Nature Human Behaviour*, 1–13. <https://doi.org/10.1038/s41562-024-01959-9>
- Chiang, C.-W., Lu, Z., Li, Z., & Yin, M. (2024). Enhancing AI-Assisted Group Decision Making through LLM-Powered Devil’s Advocate. *Proceedings of the 29th International Conference on Intelligent User Interfaces*, 103–119. <https://doi.org/10.1145/3640543.3645199>
- Chuang, Y.-S., Harlalka, N., Suresh, S., Goyal, A., Hawkins, R., Yang, S., Shah, D., Hu, J., & Rogers, T. T. (2024). *The Wisdom of Partisan Crowds: Comparing Collective*

Intelligence in Humans and LLM-based Agents.

- Collins, K. M., Sucholutsky, I., Bhatt, U., Chandra, K., Wong, L., Lee, M., Zhang, C. E., Zhi-Xuan, T., Ho, M., Mansinghka, V., Weller, A., Tenenbaum, J. B., & Griffiths, T. L. (2024). Building machines that learn and think with people. *Nature Human Behaviour*, 8(10), 1851–1863. <https://doi.org/10.1038/s41562-024-01991-9>
- Cui, H., & Yasserli, T. (2024). AI-enhanced collective intelligence. *Patterns*, 5(11), 101074. <https://doi.org/10.1016/j.patter.2024.101074>
- Gao, H., & Zhang, Y. (2024). *Memory Sharing for Large Language Model based Agents* (arXiv:2404.09982). arXiv. <https://arxiv.org/abs/2404.09982>
- Koehl, D., & Vangsnæs, L. (2023). Measuring Latent Trust Patterns in Large Language Models in the Context of Human-AI Teaming. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 67. <https://doi.org/10.1177/21695067231192869>
- Kumar, A., Tham, R.-H. M., & Steyvers, M. (2024). *Assessing the Impact of Differing Perspectives in Advice-Taking Behavior*. <https://doi.org/10.31234/osf.io/seqjr>
- Lai, V., Chen, C., Smith-Renner, A., Liao, Q. V., & Tan, C. (2023). Towards a Science of Human-AI Decision Making: An Overview of Design Space in Empirical Human-Subject Studies. *2023 ACM Conference on Fairness, Accountability, and Transparency*, 1369–1385. <https://doi.org/10.1145/3593013.3594087>
- Lu, Z., Amin Mahmo, S. H., Li, Z., & Yin, M. (2024). Mix and Match: Characterizing Heterogeneous Human Behavior in AI-assisted Decision Making. *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, 12, 95–104. <https://doi.org/10.1609/hcomp.v12i1.31604>
- Ma, S., Chen, Q., Wang, X., Zheng, C., Peng, Z., Yin, M., & Ma, X. (2024). *Towards Human-AI Deliberation: Design and Evaluation of LLM-Empowered Deliberative AI for AI-Assisted Decision-Making* (arXiv:2403.16812). arXiv. <https://arxiv.org/abs/2403.16812>
- Marjeh, R., Gokhale, A., Bullo, F., & Griffiths, T. L. (2024). *Task Allocation in Teams as a Multi-Armed Bandit*.
- McNeese, N. J., Flathmann, C., O'Neill, T. A., & Salas, E. (2023). Stepping out of the shadow of human-human team-

- ing: Crafting a unique identity for human-autonomy teams. *Computers in Human Behavior*, 148, 107874. <https://doi.org/10.1016/j.chb.2023.107874>
- Narayanan, S., Yu, G., Ho, C.-J., & Yin, M. (2023). How does Value Similarity affect Human Reliance in AI-Assisted Ethical Decision Making? *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*, 49–57. <https://doi.org/10.1145/3600211.3604709>
- Nishida, Y., Shimojo, S., & Hayashi, Y. (2024). Conversational Agent Dynamics with Minority Opinion and Cognitive Conflict in Small-Group Decision-Making. *Japanese Psychological Research*. <https://doi.org/10.1111/jpr.12552>
- Radivojevic, K., Clark, N., & Brenner, P. (2024). LLMs Among Us: Generative AI Participating in Digital Discourse. *Proceedings of the AAAI Symposium Series*, 3(1), 209–218. <https://doi.org/10.1609/aaais.v3i1.31202>
- Rastogi, C., Leqi, L., Holstein, K., & Heidari, H. (2023). A Taxonomy of Human and ML Strengths in Decision-Making to Investigate Human-ML Complementarity. *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, 11, 127–139. <https://doi.org/10.1609/hcomp.v11i1.27554>
- Sidji, M., Smith, W., & Rogerson, M. J. (2024). Human-AI Collaboration in Cooperative Games: A Study of Playing Codenames with an LLM Assistant. *Proc. ACM Hum.-Comput. Interact.*, 8(CHI PLAY), 316:1–316:25. <https://doi.org/10.1145/3677081>
- Stadler, M., Bannert, M., & Sailer, M. (2024). Cognitive ease at a cost: LLMs reduce mental effort but compromise depth in student scientific inquiry. *Computers in Human Behavior*, 160, 108386. <https://doi.org/10.1016/j.chb.2024.108386>
- Steyvers, M., & Kumar, A. (2024). Three Challenges for AI-Assisted Decision-Making. *Perspectives on Psychological Science*, 19(5), 722–734. <https://doi.org/10.1177/17456916231181102>
- Tessler, M. H., Bakker, M. A., Jarrett, D., Sheahan, H., Chadwick, M. J., Koster, R., Evans, G., Campbell-Gillingham, L., Collins, T., Parkes, D. C., Botvinick, M., & Summerfield, C. (2024). AI can help humans find common ground in democratic deliberation. *Science*, 386(6719), eadq2852. <https://doi.org/10.1126/science.adq2852>

- Tsirtsis, S., Rodriguez, M. G., & Gerstenberg, T. (2024). *Towards a computational model of responsibility judgments in sequential human-AI collaboration*. <https://doi.org/10.31234/osf.io/m4yad>
- Wegner, D. M. (1987). Transactive Memory: A Contemporary Analysis of the Group Mind. In B. Mullen & G. R. Goethals (Eds.), *Theories of Group Behavior* (pp. 185–208). Springer. https://doi.org/10.1007/978-1-4612-4634-3_9
- Westphal, M., Vössing, M., Satzger, G., Yom-Tov, G. B., & Rafaeli, A. (2023). Decision control and explanations in human-AI collaboration: Improving user perceptions and compliance. *Computers in Human Behavior*, 144, 107714. <https://doi.org/10.1016/j.chb.2023.107714>
- Yang, Z., Xu, X., Yao, B., Rogers, E., Zhang, S., Intille, S., Shara, N., Gao, G. G., & Wang, D. (2024). Talk2Care: An LLM-based Voice Assistant for Communication between Healthcare Providers and Older Adults. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 8(2), 1–35. <https://doi.org/10.1145/3659625>
- Zhu, J.-Q., Yan, H., & Griffiths, T. L. (2024). *Language Models Trained to do Arithmetic Predict Human Risky and Intertemporal Choice* (arXiv:2405.19313). arXiv. <https://arxiv.org/abs/2405.19313>