AI and Group Decision Making: An Information Processing Perspective

Thomas E. Gorman[1,2] and Torsten Reimer[3]

[1]Communication and Cognition Lab, Purdue University, USA
[2]College of Liberal Arts Research Academy
[3]Communication and Cognition Lab, Purdue University, USA

Author note

Thomas E. Gorman, https://orcid.org/0000-0001-5366-5442
Torsten Reimer, https://orcid.org/0000-0002-7419-0076

AI and Group Decision Making: An Information Processing Perspective

Introduction

Artificial Intelligence (AI) is becoming an increasingly integral component of group decision-making processes across a range of domains. From healthcare to finance, education to policymaking, AI systems are increasingly integrated into group decision-making processes, offering new avenues for enhancing efficiency, accuracy, and innovation (BaniHani et al., 2024; Burton et al., 2024; Carter & Wynne, 2024). This growing collaboration between humans and AI brings forth both significant opportunities and pressing challenges. On one hand, AI systems offer the potential to enhance information processing efficiency, improve decision accuracy, and streamline communication within teams. On the other hand, the complexities inherent in human-AI interactions—such as issues of trust and over-reliance, susceptibility to cognitive biases, erosion of critical thinking skills, lack of transparency in AI algorithms, and ethical concerns regarding accountability and fairness.

The use of AI in group settings has evolved from basic decision-support tools to more sophisticated roles, such as collaborative partners capable of generating novel insights. Large language models (LLMs), for instance, can facilitate collective intelligence by synthesizing information, generating alternative solutions, and even mediating group discussions. However, the extent to which AI enhances group performance remains context-dependent. Recent meta-analyses reveal that human-AI collaboration can lead to either augmentation of individual performance or to performance decrements (Vaccaro et al., 2024), depending on the task and interaction design.

To navigate these complexities, this chapter adopts the information processing framework as a lens for examining AI-assisted group decision-making (Hinsz et al., 1997). This framework provides a structured method to analyze how AI systems interact with human cognitive processes at each stage of decision-making. By dissecting the inputs (information acquisition and sharing), the processing mechanisms (interpretation and integration of information), and the outputs (decisions and actions), we can gain insights into the opportunities and challenges presented by AI integration.

Key questions we will seek to address within this framework::

- **Inputs**: How does AI influence the way groups search for, gather, and share information? For example, AI can augment information search through advanced data retrieval but may also introduce biases based on the algorithms' training data.
- **Processing**: In what ways do AI systems affect the interpretation and integration of information within the group? AI can facilitate complex data analysis but might obscure the reasoning process through opaque algorithms, impacting the group's shared understanding.
- **Outputs**: How do AI recommendations influence the group's final decisions and actions? The reliance on AI outputs raises questions about trust, accountability, and the potential diminishment of human agency.

This analysis is particularly timely given the rapid advancement of AI capabilities....

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. Marjieh 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.

Building upon this, McNeese et al. (2023) argue that human-autonomy teams (HATs) should be recognized as distinct from traditional human teams. They emphasize that HATs should not strive to replicate human-human team dynamics but instead should leverage the unique capabilities of AI agents. The authors propose several research trajectories to advance our understanding of HATs, including exploring diverse teaming models, redefining roles for AI teammates, expanding communication modalities, focusing on AI behavior design, developing specialized training, and emphasizing teamwork in AI design. These insights highlight the necessity of adjusting our approaches to team composition and role assignment when AI agents are involved, ensuring that both human and AI strengths are optimized in the decision-making process.

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 reccomendation 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 reccomendations 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.

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

Information Processing

Information Search

The information search stage of decision-making, once reliant on human capacity to locate and synthesize data, has been transformed by the advent of artificial intelligence (AI), particularly Large Language Models (LLMs). This section explores how AI reshapes information search, augmenting both data retrieval and synthesis, and fostering idea generation and creative discovery.

AI-Assisted Data Retrieval and Synthesis. LLMs significantly enhance the efficiency and comprehensiveness of information gathering, enabling access to a broader knowledge base and deeper insights (Bouschery et al., 2023). These models process vast datasets, identifying connections and patterns beyond human capacity. Furthermore, individual differences, such as computational thinking skills, influence how users interact with LLMs, with those possessing higher creativity and algorithmic thinking more effectively leveraging AI-generated content for deeper engagement within a specific information landscape (Flores et al., 2024). Programmers, for example, navigate between traditional web search and generative AI tools, strategically selecting between them based on factors like task familiarity and goal clarity, demonstrating the synergistic use of both resources (Yen et al., 2024). DiscipLink, for instance, uses LLMs to generate exploratory questions across disciplines, automatically expand queries with field-specific terminology, and extract themes from retrieved papers, effectively bridging knowledge gaps in interdisciplinary research (Zheng et al., 2024). Moreover, AI facilitates advanced techniques like retrieval-augmented generation (RAG), allowing LLMs to access and process real-time information, enhancing the accuracy and relevance of their output (Si et al., 2024; Wang et al., 2024). This capability empowers decision-makers with synthesized insights from diverse sources, crucial for informed choices across various fields, from scientific research to policy analysis (Burton et al., 2024).

LLM-based search tools offer natural language interfaces, streamlining complex queries and providing detailed responses, often leading to increased efficiency and user satisfaction (Spatharioti et al., 2023). However, this ease of use can also lead to overreliance on potentially inaccurate information and decreased critical evaluation, particularly when presented conversationally (Anderl et al., 2024). This can contribute to confirmation bias and the formation of "generative echo chambers," limiting exposure to diverse perspectives (Sharma et al., 2024). Furthermore, while LLMs can reduce cognitive load during information seeking, this may come at the cost of deeper learning and engagement with the material, leading to less sophisticated reasoning and argumentation (Stadler et al., 2024). Therefore, careful design and implementation are crucial to mitigate these risks and leverage the full potential of LLMs for enhanced information retrieval and synthesis.

AI in Idea Generation and Creative Discovery. AI's role extends beyond data retrieval to fostering creative discovery. LLMs act as catalysts, offering alternative perspectives, challenging assumptions, and proposing unexpected connections (Bouschery et al., 2023). In structured tasks like semantic search, AI agents enhance group performance by selectively sharing information, amplifying collective intelligence (Ueshima & Takikawa, 2024). Studies comparing human and AI-generated ideas reveal a nuanced picture: while LLMs excel at generating ideas with higher average quality (e.g., purchase intent) and even surpassing human experts in novelty (Joosten et al., 2024; Meincke et al., 2024; Si et al., 2024), they may exhibit lower feasibility (Joosten et al., 2024) and reduced diversity (Meincke et al., 2024). This highlights the importance of strategic prompt engineering, as demonstrated by Boussioux et al. (2024), who found that human-guided prompts—specifically differentiated search (i.e., prompts designed to encourage diverse and varied responses) — enhanced the novelty of LLM-generated solutions while maintaining high value.

The type of AI interaction also significantly influences human creativity. Ashkinaze et al. (2024) found that exposure to AI-generated ideas increased the diversity of collective ideas without affecting individual creativity. In contrast, H. Kumar et al. (2024) observed that while providing direct answers had minimal negative impact, exposure to LLM-generated strategies decreased both originality and creative flexibility in subsequent unassisted tasks.

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; Yan et al., 2021). 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, divded 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.

Bastola et al. (2024) further explored the potential of AI-mediated communication by examining how an LLM-based Smart Reply (LSR) system could impact collaborative performance in professional settings. They developed a system utilizing ChatGPT to generate context-aware, personalized responses during workplace interactions, aiming to reduce the cognitive effort required for message composition in multitasking scenarios. In their study, participants engaged in a cognitively demanding Dual N-back task while managing scheduling activities via Google Calendar and responding to simulated co-workers on Slack. The findings indicated that the use of the LSR system not only improved work performance—evidenced by higher accuracy in the N-back task—but also increased messaging efficiency and reduced cognitive load, as participants could more readily focus on primary tasks without the distraction of composing responses. However, it is important to note that participants expressed concerns about the appropriateness and accuracy of AI-generated messages, as well as issues related to trust and privacy. Thus, while AI-mediated communication tools like the LSR system may facilitate information sharing and alleviate cognitive demands in collaborative work, these benefits must be balanced against potential user experience challenges to fully realize their potential advantages.

- (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 Buçinca 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)

References

Anderl, C., Klein, S. H., Sarigül, B., Schneider, F. M., Han, J., Fiedler, P. L., & Utz, S. (2024). Conversational presentation mode increases credibility judgements during information search with ChatGPT. *Scientific Reports*, *14*(1), 17127. https://doi.org/10.1038/s41598-024-67829-6

Ashkinaze, J., Mendelsohn, J., Qiwei, L., Budak, C., & Gilbert, E. (2024). *How AI Ideas Affect the Creativity, Diversity, and Evolution of Human Ideas: Evidence From a Large, Dynamic Experiment* (arXiv:2401.13481). arXiv. https://doi.org/10.48550/arXiv.2401.13481

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

BaniHani, I., Alawadi, S., & Elmrayyan, N. (2024). AI and the decision-making process: A literature review in healthcare, financial, and technology sectors. *Journal of Decision Systems*, 1–11. https://doi.org/10.1080/12460125.2024.2349425

Bastola, A., Wang, H., Hembree, J., Yadav, P., Gong, Z., Dixon, E., Razi, A., & McNeese, N. (2024). *LLM-based Smart Reply (LSR): Enhancing Collaborative Performance with ChatGPT-mediated Smart Reply System* (arXiv:2306.11980). arXiv. https://arxiv.org/abs/2306.11980

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

Bouschery, S. G., Blazevic, V., & Piller, F. T. (2023). Augmenting human innovation teams with artificial intelligence: Exploring transformer-based language models. *Journal of Product Innovation Management*, *40*(2), 139–153. https://doi.org/10.1111/jpim.12656

Boussioux, L., Lane, J. N., Zhang, M., Jacimovic, V., & Lakhani, K. R. (2024). The Crowdless Future? Generative AI and Creative Problem-Solving. *Organization Science*, *35*(5), 1589–1607. https://doi.org/10.1287/orsc.2023.18430

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

Carter, W., & Wynne, K. T. (2024). Integrating artificial intelligence into team decision-making: Toward a theory of AI–human team effectiveness. *European Management Review*. https://doi.org/10.1111/emre.12685

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., & Yasseri, T. (2024). AI-enhanced collective intelligence. *Patterns*, *5*(11), 101074. https://doi.org/10.1016/j.patter.2024.101074

Flores, P., Rong, G., & Cowley, B. (2024). Information foraging in human-ChatGPT interactions: Factors of computational thinking dissociate exploration and exploitation. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *46*.

Hinsz, V. B., Tindale, R. S., & Vollrath, D. A. (1997). The emerging conceptualization of groups as information processors. *Psychological Bulletin*, *121*(1), 43–64. https://doi.org/10.1037/0033-2909.121.1.43

Joosten, J., Bilgram, V., Hahn, A., & Totzek, D. (2024). Comparing the Ideation Quality of Humans With Generative Artificial Intelligence. *IEEE Engineering Management Review*, *52*(2), 153–164. https://doi.org/10.1109/EMR.2024.3353338

Koehl, D., & Vangsness, 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

Kumar, H., Vincentius, J., Jordan, E., & Anderson, A. (2024). *Human Creativity in the Age of LLMs: Randomized Experiments on Divergent and Convergent Thinking* (arXiv:2410.03703). arXiv. https://arxiv.org/abs/2410.03703

Lu, Z., Amin Mahmoo, 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

Marjieh, 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 teaming: Crafting a unique identity for human-autonomy teams. *Computers in Human Behavior*, *148*, 107874. https://doi.org/10.1016/j.chb.2023.107874

Meincke, L., Girotra, K., Nave, G., Terwiesch, C., & Ulrich, K. T. (2024). *Using Large Language Models for Idea Generation in Innovation.* https://doi.org/Meincke, Lennart and Girotra, Karan and Nave, Gideon and Terwiesch, Christian and Ulrich, Karl T., Using Large Language Models for Idea Generation in Innovation (September 07, 2024). The Wharton School Research Paper Forthcoming, Available at SSRN: https://ssrn.com/abstract=4526071 or http://dx.doi.org/10.2139/ssrn.4526071

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/aaaiss.v3i1.31202

Sharma, N., Liao, Q. V., & Xiao, Z. (2024). Generative Echo Chamber? Effect of LLM-Powered Search Systems on Diverse Information Seeking. *Proceedings of the CHI Conference on Human Factors in Computing Systems*, 1–17. https://doi.org/10.1145/3613904.3642459

Si, C., Yang, D., & Hashimoto, T. (2024). *Can LLMs Generate Novel Research Ideas? A Large-Scale Human Study with 100+ NLP Researchers* (arXiv:2409.04109). arXiv. https://doi.org/10.48550/arXiv.2409.04109

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

Spatharioti, S. E., Rothschild, D. M., Goldstein, D. G., & Hofman, J. M. (2023). *Comparing Traditional and LLM-based Search for Consumer Choice: A Randomized Experiment* (arXiv:2307.03744). arXiv. https://doi.org/10.48550/arXiv.2307.03744

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

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

Ueshima, A., & Takikawa, H. (2024). Discovering Novel Social Preferences Using Simple Artificial Neural Networks. *Collabra: Psychology*, *10*(1), 121234. https://doi.org/10.1525/collabra.121234

Vaccaro, M., Almaatouq, A., & Malone, T. (2024). When combinations of humans and AI are useful: A systematic review and meta-analysis. *Nature Human Behaviour*, 1–11. https://doi.org/10.1038/s41562-024-02024-1

Wang, R., Zhou, X., Qiu, L., Chang, J. C., Bragg, J., & Zhang, A. X. (2024). *Social-RAG: Retrieving from Group Interactions to Socially Ground Proactive AI Generation to Group Preferences* (arXiv:2411.02353). arXiv. https://doi.org/10.48550/arXiv.2411.02353

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

Yan, B., Hollingshead, A. B., Alexander, K. S., Cruz, I., & Shaikh, S. J. (2021). Communication in Transactive Memory Systems: A Review and Multidimensional Network Perspective. *Small Group Research*, *52*(1), 3–32. https://doi.org/10.1177/1046496420967764

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

Yen, R., Sultanum, N., & Zhao, J. (2024). To Search or To Gen? Exploring the Synergy between Generative AI and Web Search in Programming. *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, 1–8. https://doi.org/10.1145/3613905.3650867

Zheng, C., Zhang, Y., Huang, Z., Shi, C., Xu, M., & Ma, X. (2024). DiscipLink: Unfolding Interdisciplinary Information Seeking Process via Human-AI Co-Exploration. *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology*, 1–20. https://doi.org/10.1145/3654777.3676366

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