Designing Support for Human-AI Idea Selection: Human Agency and AI Autonomy

MEREDITH YOUNG-NG, University of California, Davis, USA
QINGXIAOYANG ZHU, University of California, Davis, USA
JINGXIAN LIAO, University of California, Davis, USA
HAO-CHUAN WANG, University of California, Davis, USA

Many human-AI co-creation systems focus on leveraging generative AI's capabilities to improve creative content generation. However, AI support in creative idea selection remains underexplored. We used a technology probe (IdeaMap) with an AI collaborator (gpt-3.5-turbo) in a pilot study to elicit user perceptions of human agency and AI autonomy on idea selection. We share our findings and discuss how we can design human-AI co-creative systems that support co-selection through a balance of human agency and AI autonomy.

Additional Key Words and Phrases: human-AI collaboration, human agency, AI autonomy, ideation maps

1 Introduction

Technical advancements in generative AI, especially large language models (LLMs), have shifted the paradigm of how people think of AI, and interact with AI systems, from "AI as tool" to "AI as partner" and to "AI as agent" [8, 15]. This paradigm-shift is mainly driven by technological optimism: a general belief of what generative AI as a technology can potentially do now, but not a grounded understanding of what actual users would and would not use the technology for, as well as to what extent those users would be willing to trust, share control, and closely collaborate with AI on critical tasks. While there have been some initial demonstrations on how an AI co-worker can be conceptualized (and to some extent operationalized), such as a co-piloting chatbot [13] or a joint decision maker [11], development and evaluation focuses mostly on output correctness or similarity to human performance. There remains a lack of deeper inquiries looking into the socio-cognitive processes of how people and AI can work as co-workers of a team on specific group tasks, and what interactive mechanisms and collaboration support would benefit the functioning of human-AI teams.

Group brainstorming is a common group task central to different types of knowledge work ranging from decision making to creative design. In a conventional design thinking process, group brainstorming has two cyclical phases of socio-cognitive work: divergent thinking (idea generation: producing a large quantity of ideas based on a given problem) followed by convergent thinking (idea selection: filtering the generated ideas to a smaller set of ideas based on their quality and relevance to the problem) [4, 5]. During group brainstorming, it is critical that team members fully utilize and integrate individual cognitive resources, such as knowledge, experience, and reasoning capacities, to overcome creativity barriers that workers may face individually. Now that brainstorming groups can consist of human and AI co-workers, there's a need to revisit the design space of group brainstorming support to understand how AI participation can impact different aspects of group brainstorming and more generally human-AI co-creation systems.

Authors' Contact Information: Meredith Young-Ng, mjyoungng@ucdavis.edu, University of California, Davis, Davis, California, USA; Qingxiaoyang Zhu, qinzhu@ucdavis.edu, University of California, Davis, Davis, California, USA; Jingxian Liao, jxliao2021@gmail.com, University of California, Davis, Davis, California, USA; Hao-Chuan Wang, hciwang@ucdavis.edu, University of California, Davis, Davis, California, USA.

Many existing human-AI co-creation systems focus on using AI's generative capabilities to boost creative content generation in either textual [20] and/or image formats [7]. Having a LLM brainstorm with people to produce useful ideas for knowledge work requires a careful scrutiny of human agency (what humans can do socially) and AI autonomy (what AIs can do technically), and an integration of the two [16]. Idea generation in human-AI group brainstorming has gotten more attention due to generative AI's unique information association, reassembly, and reproduction capacities [9]. However, idea selection in group brainstorming, commonly known as the phase where co-workers have to converge their thoughts and assess idea candidates to identify valuable ones, tends to be overlooked in the current demonstrations of human-AI co-creation: users are left to curate the final results alone without deliberate design to work with AI.

Designing for human-AI idea selection is complicated: humans lack the same human-human communication paradigms with AI teammates. Simpler conversation-level interactions and outeractions [17] are missing: each human utterance is translated into textual input as part of an LLM prompt, which uses an API to transform the input into parameter arguments, generate a response, and relay it back to the humans. Furthermore, idea selection in groups is difficult: it requires group convergence on a "final" idea. In human-only teams, individuals rarely instinctively align on the same idea due to reasons such as self-interest, ownership over ideas, and team power dynamics. They must communicate with each other, establishing a shared understanding and grounding their conversation until they arrive at a final decision [3]. Such conversational grounding does not yet exist in human-AI teams.

AIs are not like humans: while good at generation, AIs' motivations remain unknown. Their status is not naturally visible to humans, impacting how humans interact conversationally with them and thus human-AI idea selection. It is unclear how humans perceive AI teammates during co-selection processes, particularly how AIs' actions may be interpreted as social cues. AIs naturally lack the same set of social cues as humans: how do humans know when an AI wants to own an idea? Does AI always need to explain its rationale behind idea selection to be deemed credible or trustworthy? In human-only groups, implicit social rules in conversations support awareness and natural self-organization into a team structure. But when lacking known social cues from one another, can humans and AIs still self-organize into effective human-AI teams? Fundamentally, an idea selection's successes and failures are inherently tied to their outcome: the "final" ideas. Everyone involved in selecting an idea is directly responsible and accountable for the joint decision. Ultimately, these principles of visibility, awareness, and accountability are critical to human-AI communication and building a corpus of shared knowledge that can be used in idea selection. It also suggests the possibility of employing social translucence design principles [6] when designing for human-AI idea selection systems.

Levels of human agency and AI autonomy dynamically shift throughout human-AI group ideation [9]. In this workshop paper, we share results from a pilot study on user perceptions of human agency and AI autonomy during human-AI group brainstorming and discuss how we can design human-AI co-creative systems that support human-AI idea selection through a balance of human agency and AI autonomy. To elicit these user perceptions, 12 participants (all experienced with AI) used IdeaMap, a technology probe designed with a chatroom and ideation map that functions as a shared knowledge workspace and an AI collaborator (gpt-3.5-turbo) in a series of design workshops. We found that while participants acknowledged potential benefits of a more active AI teammate during earlier iterations in idea selection, they ultimately preferred having human control over final idea selection. For further details, please refer to our extended abstract at CHI 2025 [21]. Based on these findings, we aim to facilitate discussion on how to design generative AI co-creative systems that support idea selection, particularly with features that promote social translucence and shared ownership of ideas between humans and AIs.

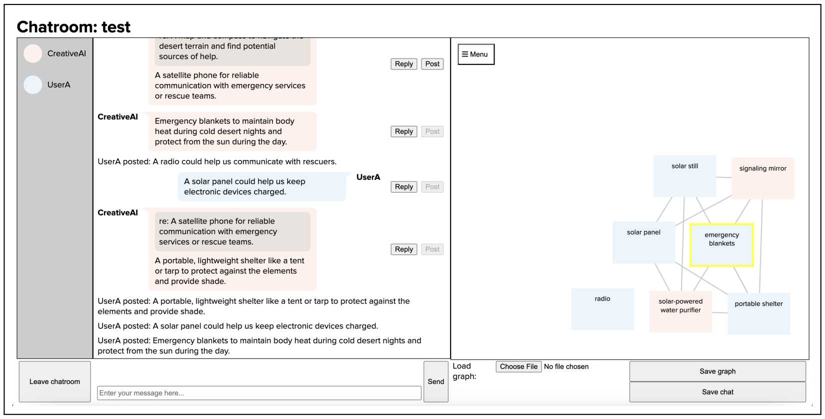

Fig. 1. IdeaMap human-AI group brainstorming system, featuring a sidebar with all usernames and their associated colors (left), a chatroom for communicating and sharing ideas (middle), and an ideation map visualization based on posted ideas (right).

2 Pilot Study

IdeaMap is designed as a technology probe [10] to explore how generative AI can act like a teammate by assisting with idea generation and selection. We include a chatroom and ideation map visualization as interaction modalities in a Flask-based web app to let levels of human agency and AI autonomy fluctuate throughout group brainstorming (Figure 1). Inspired by post-it brainstorming with digital affordances, the ideation map functions as a shared knowledge workspace throughout multiple iterations of convergent thinking. Ideas are connected, and thus clustered, based on their semantic distance to help users visualize similarities between ideas. Users can select ideas from text messages by clicking the respective "Post" button, modify ideas directly on the map in edit mode, and filter ideas on the map in delete mode. Implemented with gpt-3.5-turbo chat completions, the AI collaborator can suggest and post ideas, but not edit or delete ideas so human users retain full control over the final output, consistent with prior work on selection [2].

We recruited 12 participants (age 23-36 years old; 7 female, 5 male) from local university networks that are native or fluent in English and experienced with AI from our local university network. Each study session had 2 participants working with the creative AI collaborator. Our 1.5 hour study had 4 components: a 5-minute system tutorial, two technology probes (combined 25 minutes), a 1-hour participatory design workshop, and a 1-minute demographic questionnaire. In the technology probes, participants used IdeaMap in a single-party human-AI team (1 human, 1 AI) for 10 minutes, then in a multi-party human-AI team (2 humans, 1 AI) for 15 minutes. Topics included a desert and moon survival scenario inspired by the NASA moon survival team building exercise [18], assigned in a counterbalanced order. Participants were instructed to brainstorm new ideas in the chatroom, post "good" ideas to the ideation map, and filter posted ideas down to the best ones using the map edit and delete modes. During the semi-structured participatory design workshop, participants used a Miro board [14] for digital post-it note brainstorming, discussing which features an AI collaborator should have and the level of AI involvement in idea selection within a group brainstorming context.

We created affinity diagrams from participants' Miro boards to derive themes about participants' perceptions of human agency in AI autonomy in group brainstorming [12]. Consistent with prior work on human-AI teaming tasks [1, 2, 19], participants expressed a preference for user control and overriding AI autonomy on the final set of selected ideas. We found that a majority of participants preferred a more passive AI collaborator for idea selection in both single-party teams (P3-P9, P11-P12) and multi-party teams (P1, P3-P5, P7-P8, P11-P12). Participants cited doubts about AI's capabilities to make good idea selections (P2, P7), uncertainty on AI's trustworthiness (P8, P11), and a preference for working in multi-party (multiple human) groups to have "human verification of [AI's] idea selection" (P5, P6). While participants seem skeptical of the current state-of-the-art AI's involvement in later stages of co-selection, they acknowledge the potential of AI's help in earlier stages of idea selection. Human-AI teams could leverage AI's computational strengths for organizational efficiency by "[helping] the user find key terms" for faster iteration (P2) and automatically filtering and combining similar ideas (P1-P4), which could be useful when working in large groups with many ideas (P12). Selection speed could also be improved when an AI teammate serves as a mediator during disagreements between human teammates in multi-party teams (P7-P9, P12). AI teammates can also play other supportive roles that impact human teammates' contributions, such as asking passive teammates to contribute (P5-P6, P10, P12), selecting a human-generated idea, or contributing via a human-selected, AI-generated idea. P10 noted that when they selected an AI idea, it made them feel more confident in their decisions, as if they "were agreeing with an AI idea." A more transparent AI that explains its idea selections could further amplify this confidence boost and improve AI's trusthworthiness (P7-P8).

3 Discussion and Future Work

While participants preferred a more passive AI and human control over final human-AI idea selection, we identified possibilities for active AI support in earlier stages of co-selection. These AI interactions include awareness of human teammates' activity by individually prompting passive teammates and accountability via shared ownership of AI-contributed ideas. This motivates us to further investigate these possible features of human-AI idea selection in a more formal study to better understand how to design AI support in co-creative idea selection systems.

Designing support for human-AI idea selection can start with the interaction-level system components from social translucence [6]. Making the AI's status visible to human teammates can improve the human-AI team's social coordination. These visual signals could go beyond what ideas are ultimately "selected" by AI to indicate and explain ideas that have been considered, but not selected. Publicizing these AI thought processes could help inform human reactions, user perceptions of an AI's trustworthiness, and thus joint decisions on idea selection.

In human-AI teams, human and AI teammates need to establish a set of social rules. Human and AI teammates that are aware of these rules and attuned to each other's social cues can help the group better converge on idea selection. For example, a system design could promote a team formation of equal human and AI partners by giving humans and AIs the same set of actions (and thus opportunities for human agency and AI autonomy) within the shared knowledge workspace. Furthermore, humans and AIs have different strengths: for example, AIs' computational power enables them to "think" much faster than humans. An AI teammate aware of humans' computational speed, like the creative AI in IdeaMap, could slow down its actions to match the pace of humans' interactions. This could foster greater shared knowledge between human and AI teammates which could better inform the team's idea selection. AI teammates could also be aware of human social cues indicating a lack of activity and respond by prompting inactive users to participate. In parallel, documentation of human and AI actions can help provide accountability for contributions to idea selection. Visualizing these actions, such as displaying which teammates selected an idea, can help promote shared ownership and thus responsibility for ideas. This accountability could help enforce the team's social rules for collaboration.

We hope to get feedback on our research direction and approaches for future experimental designs in this workshop. The AI design in our current technology probe, IdeaMap, supports some awareness of human social cues, but has limitations such as a lack of visible AI status. Specifically, we hope to elicit insights from workshop participants on what design features for supporting human-AI idea selection should be explored: how can we design an AI that supports idea convergence, but also retains human agency including human dissent? We would also like to brainstorm strategies for fostering grounding in communication between human and AI teammates. Participants in our pilot study indicated a preference for greater human agency, particularly over the final idea results. Given this desire for greater human control, can humans and AIs agree that humans and AIs are equal partners in human-AI idea selection tasks? Humans often exercise their agency when prompting the AI to take an action, but are they equally receptive to AI instructions or guidance? Finally, we would like to discuss what AI design features promote shared ownership. Humans and AIs have accountability based on whether they are actively contributing to the shared knowledge space and the idea selection process. How can we design social cues for AIs, such as a cue that expresses an AI's desire for ownership or a cue that recognizes a human need for further AI explanation? We hope to share our experiences and insights with designing human-AI collaboration and co-creative systems and other related topics in the workshop.

References

[1] Saleema Amershi, Dan Weld, Mihaela Vorvoreanu, Adam Fourney, Besmira Nushi, Penny Collisson, Jina Suh, Shamsi Iqbal, Paul N Bennett, Kori Inkpen, et al. 2019. Guidelines for human-AI interaction. In *Proceedings of the 2019 chi conference on human factors in computing systems*. 1–13.

[2] Onur Asan, Alparslan Emrah Bayrak, Avishek Choudhury, et al. 2020. Artificial intelligence and human trust in healthcare: focus on clinicians. *Journal of medical Internet research* 22, 6 (2020), e15154.

[3] Herbert Clark and Susan Brennan. 1991. Grounding in communication. *Perspectives on socially shared cognition* (1991).

[4] Design Council. 2025. The Double Diamond. https://www.designcouncil.org.uk/our-resources/the-double-diamond/

[5] Rikke Friis Dam. 2025. 5 Stages in the Design Thinking Process. https://www.designcouncil.org.uk/our-resources/the-double-diamond/

[6] Thomas Erickson and Wendy A Kellogg. 2000. Social translucence: an approach to designing systems that support social processes. *ACM transactions on computer-human interaction (TOCHI)* 7, 1 (2000), 59–83.

[7] Xianzhe Fan, Zihan Wu, Chun Yu, Fenggui Rao, Weinan Shi, and Teng Tu. 2024. ContextCam: Bridging Context Awareness with Creative Human-AI Image Co-Creation. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–17.

[8] Jonathan Grudin. 2017. *From Tool to Partner: The Evolution of Human-Computer Interaction* (1 ed.). Springer Cham.

[9] Jessica He, Stephanie Houde, Gabriel E Gonzalez, Darío Andrés Silva Moran, Steven I Ross, Michael Muller, and Justin D Weisz. 2024. AI and the Future of Collaborative Work: Group Ideation with an LLM in a Virtual Canvas. In *Proceedings of the 3rd Annual Meeting of the Symposium on Human-Computer Interaction for Work*. 1–14.

[10] Hilary Hutchinson, Wendy Mackay, Bo Westerlund, Benjamin B Bederson, Allison Druin, Catherine Plaisant, Michel Beaudouin-Lafon, Stéphane Conversy, Helen Evans, Heiko Hansen, et al. 2003. Technology probes: inspiring design for and with families. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 17–24.

[11] 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*. 1369–1385.

[12] Andrés Lucero. 2015. Using affinity diagrams to evaluate interactive prototypes. In *Human-Computer Interaction–INTERACT 2015: 15th IFIP TC 13 International Conference, Bamberg, Germany, September 14-18, 2015, Proceedings, Part II 15*. Springer, 231–248.

[13] Microsoft. 2025. Microsoft Copilot. https://copilot.microsoft.com/

[14] Miro. 2025. Miro | The Innovation Workspace. https://www.miro.com

[15] Caterina Moruzzi and Solange Margarido. 2024. A user-centered framework for human-ai co-creativity. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. 1–9.

[16] Michael Muller and Justin Weisz. 2022. Extending a human-ai collaboration framework with dynamism and sociality. In *Proceedings of the 1st Annual Meeting of the Symposium on Human-Computer Interaction for Work*. 1–12.

[17] Bonnie A Nardi, Steve Whittaker, and Erin Bradner. 2000. Interaction and outeraction: Instant messaging in action. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work*. 79–88.

[18] NASA. 2016. NASA Exercise: Ranking Survival Objects for the Moon. https://www.psychologicalscience.org/observer/nasa-exercise

[19] Ben Shneiderman and Catherine Plaisant. 2010. *Designing the user interface: strategies for effective human-computer interaction*. Pearson Education India.

[20] Sangho Suh, Meng Chen, Bryan Min, Toby Jia-Jun Li, and Haijun Xia. 2024. Luminate: Structured Generation and Exploration of Design Space with Large Language Models for Human-AI Co-Creation. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–26.

[21] Meredith Young-Ng, Qingxiaoyang Zhu, Jingxian Liao, and Hao-Chuan Wang. 2025. Balancing Human Agency and AI Autonomy in Human-AI Idea Selection. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. 1–6.