Transitioning Focus: Viewing Human-AI Collaboration as Mixed-focus Collaboration

Zhuoyi Cheng
College of Software Technology
Zhejiang University
Hangzhou, China
zy_cheng@zju.edu.cn

Pei Chen*
College of Computer Science and Technology
Zhejiang University
Hangzhou, China
chenpei@zju.edu.cn

Yiwen Ren
College of Computer Science and Technology
Zhejiang University
Hangzhou, China
renyiwen@zju.edu.cn

Wenzheng Song
College of Computer Science and Technology
Zhejiang University
Hangzhou, China
svvz@zju.edu.cn

Lingyun Sun
College of Computer Science and Technology
Zhejiang University
Zhejiang – Singapore Innovation and AI Joint Research
Hangzhou, China
sunly@zju.edu.cn

Abstract

The abilities that recent AI models presented, like multi-modal content generation and reasoning, allow us to see the possibility of human-AI collaboration. However, enabling AI to act proactively and harmoniously in collaboration still faces challenges, and aspects like the optimal action timing and collaboration dynamism, await to be explored. Exploring these aspects is important to designing adaptive AI to enhance the human-AI collaboration experience and system usability. In this study, we cut in from the collaboration level and view human-AI collaboration as mixed-focus collaboration to focus on human's transitions between independent and collaborative works. Grounded on previous studies in human-human collaboration, we identified four coupling styles and seven types of transition cues in human-AI collaboration, serving as preliminary results for future studies. We envisioned how our results could be further extended to support the design of adaptive AI, hoping to enhance human-AI collaboration experience and the usability of collaborative systems.

CCS Concepts

• **Human-centered computing → Collaborative interaction**; **Collaborative and social computing theory, concepts and paradigms**.

Keywords

Human-AI Collaboration, Mixed-focus Collaboration, Coupling Style, Adaptive AI, Design Collaboration

ACM Reference Format:
Zhuoyi Cheng, Pei Chen, Yiwen Ren, Wenzheng Song, and Lingyun Sun. 2025. Transitioning Focus: Viewing Human-AI Collaboration as Mixed-focus Collaboration. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25), April 26-May 1, 2025, Yokohama, Japan.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3706599.3720109

1 Introduction

Recent large-scale models allow Artificial Intelligence (AI) to not only generate and reason multi-modal content in real-time [8, 32], but also analyze the context to participate in work proactively [18, 20, 22, 25]. These abilities make AI being a collaborator in people's work prospective. However, enabling such AI agents to collaborate with humans harmoniously still faces challenges for bringing possible interruptions and negatively affect human perceptions toward AI [17, 19]. Therefore, researchers begin to study the basic aspects and descriptive frameworks of such collaboration to enhance user experience and the usability of human-AI collaborative systems. For example, Kuang et al. [17] explored the optimal timing for AI to detect and propose usability problems, and He et al. [12] modeled the dynamism of human and AI initiative during a group brainstorming session to instruct future integration of AI in group work.

Breaking down the collaboration process and looking into detailed aspects facilitate the development of adaptive AI to fit in the dynamically changing collaboration process, where people's performance and knowledge [3, 27, 28], task types and goals [22], collaboration patterns [14, 23], etc., all change as collaboration proceed [34]. Existing attempts have implemented adaptive AI autonomy and gained higher team performance in a shared workspace setting [25], adaptive communication strategies to cater to users' changing needs and expectations [16, 18], adaptive modeling of user properties to provide personalized interaction [4], etc. In this study, we focus on the collaboration patterns in human-AI collaboration,

*Corresponding author

CHI EA '25, April 26-May 1, 2025, Yokohama, Japan

ACM ISBN 979-8-4007-1395-8/2025/04
https://doi.org/10.1145/3706599.3720109

Table 1: Dimensions used to categorize and describe coupling styles in different studies.

	Focused Problem	Level of Engagement	Engaged Way	Engaged Area	Subgroup	Group Distribution
Tang et al. [30]	✓	✓	✓	✓		
Tuddenham and Robinson [31]	✓	✓	✓	✓		
Isenberg et al. [14]	✓	✓	✓	✓		
Brudy et al. [1]	✓	✓	✓	✓	✓	
Neumayr et al. [23]	✓	✓	✓	✓	✓	✓
Saffo et al. [24]	✓	✓	✓	✓		

aiming at adding insights into future design of adaptive AI that can adjust its behaviors in the changing and evolving collaboration process. Inspired by previous research in computer-supported cooperative work, we view human-AI collaboration as mixed-focus collaboration, specifically studying how humans transition between independent and collaborative works when collaborating with AI, and envision how AI can cater to such transition.

Human-human collaboration often involves mixed-focus collaboration, where people transition occasionally between independent and collaborative works [10, 30]. The collaboration patterns are thus dynamic during the process, presenting different coupling styles with divergent behavioral characteristics [14, 23, 30]. We observed similar phenomena in human-AI collaborations, where humans sometimes actively collaborated and discussed with AI, and at other times worked independently and were inclined to ignore AI. These behavioral differences are often accompanied by conflict requirements in independent and collaborative works with separate needs for operations, communications, collaboration, etc. [10, 29, 30]. When working independently, people tend to focus on interacting with the workspace and their own tasks, while collaborative works require communications and coordination between collaborators to solve the same problem [10, 14, 23]. Such phenomena imply a need for AI adaptivity to fit in humans' changing behaviors and prompt fluid coupling transition, which is considered necessary for collaborative systems to alleviate collaboration friction [15, 21, 30]. However, to the best of our knowledge, there is a dearth of studies that support these mixed-focus features in human-AI collaboration well, which can serve as one of the prerequisites for designing adaptive AI behaviors in human-AI collaborative systems.

Therefore, we view human-AI collaboration as mixed-focus collaboration in this study. Our main goal is to identify coupling styles and coupling transition cues in human-AI collaboration as a prequel to designing adaptive human-AI collaborative systems that address the challenges brought by the mixed-focus features. Similar to Muller et al. [22] and Salikutluk et al. [25], our study is contextualized in human-AI collaborative design in a shared workspace setting, whose process is complex and dynamic [35] and involves both independent and collaborative work. To achieve our goal, we first extracted the dimensions for describing coupling styles and common coupling transition cues based on previous mixed-focus collaboration studies [13, 23, 24, 30, 31] (§2), then analyzed 12 video recordings to identify the coupling styles and coupling transition cues appeared in the human-AI collaboration process (§3). The analysis generated four coupling styles and seven types of coupling transition cues (§4), which we used to envision and instruct future design of adaptive AI in human-AI collaboration (§5).

Our main contributions are three-fold: 1) we demonstrated the mixed-focus features in human-AI collaboration with empirical evidence, 2) we identified four coupling styles in human-AI collaboration and concluded their characteristics, and 3) We also envisioned how these findings could guide the future design of AI behaviors in human-AI collaborative systems that can adapt to the evolving collaboration process, hoping to contribute to more fluent and human-centric collaborative experiences, and enhancing such systems' usability.

2 Mixed-focus Collaboration and Coupling Style

2.1 Definition and Development

Humans transition between independent and collaborative works in collaboration, described as mixed-focus collaboration [10, 11]. Such transitions are driven by the level of dependency on one's collaborators, i.e., how much work people can do before collaborating with others [26]. Since the states of working independently and collaboratively are not polarized, and instead change along a continuum [23], researchers use the level of coupling to describe the overall collaborative states when the level of dependency differs [26]. That means, when people are coupled more tightly, they have a higher dependency on each other to advance work, resulting in more frequent interaction with collaborators. On the other hand, they depend less on others and interact more with the workspace and artifact when they are coupled more loosely [10, 14, 23].

Later researchers identified coupling styles to describe different levels of coupling through observational studies [30], which has gone through several iterations and updates to fit in different contexts [1, 14, 31], and support the analysis of collaboration dynamism and the design of collaborative systems [23]. During this process, the connotation of "level of dependency" has been gradually clarified and enriched, manifested by the dimensions for describing coupling styles (Table 1). In the work of Tang et al. [30] about co-located collaboration using tabletop displays, they defined six coupling styles considering the differences in collaborators' focused problems, the level of engagement, engaged areas, and engaged ways (operation, view, communication). Subsequent works mainly extend Tang et al.'s work, with some merging several styles (e.g., [31]) and others subdividing (e.g. [14]). Brudy et al. [1] applied this concept to co-located collaboration involving more than two people and devices, and considered the state of subgroups, which is also an important dimension considered by Neumayr et al. [23]. Different from all previous work, Neumayr et al. took group distribution into consideration and developed a descriptive framework for analyzing hybrid collaboration, which involves both co-located and distributed collaborators.

2.2 Coupling Transition

The original goal for identifying coupling styles was for an in-depth study of human's transitions between independent and collaborative work [30]. Although such transitions are observed to be rapid and opportunistic [5, 23, 31], researchers still identified cues that can trigger people to transition between coupling styles, like watching collaborators' gestures and body movements in co-located(-like) collaborations [30, 31]. Observing such cues and initiating coupling transition is the result of workspace awareness's function [11, 31], but in distributed collaborations where workspace awareness requires extra supporting mechanisms, coupling transition cues can become more conspicuous. For example, Grønbæk et al. [9] found people's reconfiguration of their whiteboards could signal coupling transitions when collaborating through embodied whiteboards, and checking for collaborators' views is considered to stand for transition and maintain tight coupling. Miller et al. [21] designed an online collaborative content creation system that enables users to leave visual traces when peeking at others' views to support transitions. Recent studies have leveraged AI to detect students' learning states (e.g., get stuck) to assist teachers in managing coupling transitions and orchestrating classrooms during collaborative learning [6, 33]. While humans can transition between coupling styles by attending to various cues based on their experiences, and AI can detect transition cues upon proper training, it is still unclear how can AI manage coupling transition in human-AI collaboration due to the lack of in-depth study on both the characteristics of coupling styles and the cues used to manage coupling. Therefore, this study aims to ground future design of adaptive AI that can transition fluidly between coupling styles in human-AI collaboration by giving a preliminary conclusion on coupling styles and corresponding cues when people initiate transitions.

3 Method

3.1 Analysis Materials

We used data from six participants (three females and three males) in a previous human-AI collaborative design experiment[2] for identifying coupling styles and coupling transition cues. Participants' ages ranged from 22-24 (M = 22.83, SD = .98), with at least two years of experience in design and previous experiences using generative AI. The experiment involved a human-AI collaborative design task, where participants were required to design e-scooter or headphone concepts in a collaborative system. We conducted the experiment using the Wizard-of-Oz method with a within-subject setting: two wizards worked jointly to simulate an AI collaborator working proactively with a designers to propose headphone/e-scooter concepts, and controlled *whether AI would consider designers' design activities and current working content when generating proactive feedback* (Figure 1). Since this study aims to obtain coupling styles and transition cues as many as possible, our later analysis would not consider the differences caused by the within-subject variable between the two conditions.

The data we used included 12 experiment recordings and participants' communication histories. Each recording lasted about 40 minutes (segments with errors occurred were excluded), covering situations including participants communicating with AI through speech, textual input, and mixed methods to enlarge the diversity of our coding results.

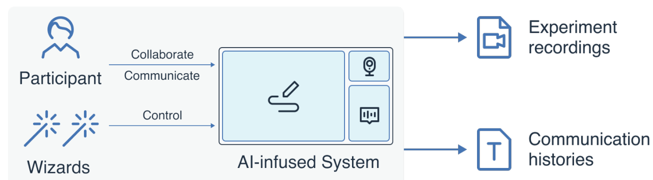

Figure 1: Experiment setting and data for analysis.

3.2 Analysis Method

Before analysis, the video recordings were first pre-processed to exclude invalid segments (i.e., talking to experimenters, errors occur, and experimenters' management of process), then subdivided into clips of 90 seconds to judge the transitions more accurately. Next, we prepared a descriptive coupling style coding scheme (see Appendix A) based on that of Neumayr et al. [23] using the following two dimensions in Table 1: focused problem and engaged way. Specifically, we distinguished between different engaged ways, involving communication and operation. The remaining four dimensions were not considered because in our experiment: 1) there was barely a difference in the level of engagement, 2) AI could not decide its engaged area, 3) only human-AI dyads were involved, and 4) human-AI collaboration is hard to be treated as either co-located or distributed collaboration.

Two of the authors analyzed the video recordings to identify coupling styles and extract transition cues. They identified the coupling styles in each 90-second clip according to the coding scheme, and were allowed to modify the starting and ending points of the clips, which were controlled to range in length of 30-120 seconds. The transition cues were open coded, and two coders aimed to extract as many diverse cues as possible. Each of them first analyzed one participant and discussed if new definitions or new coupling styles emerged until a consensus was reached. Then, the two coders each analyzed two participants, and cross-checked each other's analysis results. Two coders reached a good level of inter-coder reliability (Cohen's Kappa = .782) on coupling style identification. They later discussed discrepancies to achieve a complete agreement iteratively, leading to the final analysis results.

4 Results

4.1 Identified Coupling Styles

We identified four coupling styles in human-AI collaboration: *Discussion (DISC)*, *On Standby (OSTB)*, *Slightly Collaborated (SCOL)*,

Table 2: The analysis results of coupling styles in human-AI collaboration and corresponding descriptions.

	Focused Problem	Engaged Way	
		Communication	Operation
DISC	Same	Very frequent	Sometimes
OSTB	Same	Very infrequent	Infrequent
SCOL	Similar	Sometimes	Frequent
INDW	Related	Very infrequent	Very frequent

	DISC			OSTB	SCOL	INDW
	DISC-M	DISC-U	Σ			
P1	37.62%	26.46%	64.08%	6.55%	18.69%	10.68%
P2	11.11%	13.71%	24.82%	16.78%	58.39%	0.00%
P3	9.73%	15.21%	24.94%	2.24%	57.86%	14.96%
P4	42.28%	7.82%	50.10%	2.40%	43.69%	3.81%
P5	23.10%	16.19%	39.29%	0.00%	36.19%	24.52%
P6	7.06%	40.39%	47.45%	0.00%	38.44%	14.11%

Figure 2: The percentage of time spent on each coupling style.

and Independent Working (INDW). Among them, *DISC*, *SCOL*, and *INDW* originated from *DISC*, *SIDC*, and *SGP* in Neumayr et al. [23], respectively, for having the same levels of descriptive dimensions, but whose definitions were adapted to describe collaborative behaviors in our context better. *OSTB* was a newly observed coupling style. The definitions are detailed below. We consider the first three (identified with round parentheses) to be tightly coupled, with the level of coupling descending, and the remaining one (identified with square brackets) to be loosely coupled. Table 2 presents the coupling styles described by the aforementioned dimensions.

(DISC): The participant is actively discussing with AI and sometimes operates in the workspace, involving writing or pasting keywords, drawing simple marks, etc. We further subdivide this coupling style into *Discussion-Mutual* (***DISC-M***) and *Discussion-Unilateral* (***DISC-U***) considering the differences in whether the participant cares about AI's proactive communication.

(OSTB): The participant requires information from AI to proceed with the design process and reacts immediately after AI generates information. If the participant has asked AI a question before, (s)he usually operates in the workspace aimlessly, like dragging the canvas back and forth, scrolling the conversation history area rapidly, etc. when waiting for results In other cases, the participant stops communicating or operating for not yet deciding on what to do next, until AI communicates proactively.

(SCOL): The participant sometimes initiates communication with AI but not very often, while actively working in the workspace and is inclined to have a shufti at, refer to, or ignore AI's proactive communication.

[INDW]: The participant mainly works independently in the workspace to sketch or organize information and communicates with AI very infrequently. AI's proactive communication is usually ignored.

We also visualized the percentage of time each participant spent on each coupling style in each condition using the heat map (Figure 2). Overall, human-AI collaboration was predominantly tightly coupled (*DISC* and *SCOL*), with five participants spending less than 15% of the total time on *INDW*, which is aligned with previous studies in human-human collaboration [7, 30]. The *OSTB* style also takes up only a small proportion in most cases (five participants spent less than 7% of the total time), which we inferred to be caused by the transitional nature of *OSTB*.

4.2 Identified Cues for Coupling Transition

We recorded 88 coupling transitions and extracted 103 transition cues from the video recordings. Participants averagely made 7.33 transitions ($SD = 3.17$) in each 40-minute session. We analyzed the transition cues iteratively using thematic analysis and obtained two categories and seven codes, including four types of content-level codes and three types of behavior-level codes (Table 3). The frequency distribution of each code in different coupling style transitions is illustrated in Figure 3.

Extra information is required (EI): 19 transition cues were categorized into **EI**. Participants were prone to make coupling transitions when they needed extra information, including complementing solutions to certain design requirements, confronting problems beyond their knowledge, and requesting AI to generate images. These cues were commonly seen when transitioning to *OSTB* and *SCOL*, mostly seen in *SCOL* → *OSTB* and *INDW* → *SCOL*, where participants would simply wait for answers from AI or continue doing work in hand while attending to AI's feedback.

Concerned information appears (CI): 18 transition cues were categorized into **CI**. When AI mentions what the participants find interesting, valuable, or debatable, they might transition from one coupling style to another. These cues were commonly seen when transitioning to *DISC-M* and from *DISC-M* to *DISC-U*, and occasionally seen when transitioning to *SCOL*.

Previous goal has (not) been achieved (GA): 18 transition cues were categorized into **GA**. Cues in this code were more unpredictable because they were the end of some common sub-tasks and often appeared without heralds, like obtaining enough information, ending sketching, ending organizing information, etc. These cues were scattered in eight types of transitions, but mostly seen when transitioning from *OSTB*.

Task step switch (SS): 16 transition cues were categorized into **SS**. Since the experiment set four predefined steps, participants' coupling styles usually change when they need to proceed to the next step. Similar to **GA**, this code was also scattered and identified in ten types of transitions.

Operation in the workspace increases (OI): 19 transition cues were categorized into **OI**. When participants started to sketch product concepts, organize existing information, and connect different parts using marks, they were likely to transition to looser coupling styles. These cues were only seen when transitioning from tighter to looser coupling styles, mostly seen in *DISC-M* → *SCOL* and *SCOL* → *INDW*, where participants depend less on AI outputs and more on their own thoughts and existing information.

Operate without adding information (WI): 10 transition cues were categorized into **WI**. These cues were only seen when transitioning to *OSTB*, where participants scrolled the canvas and conversation history area rapidly, modified textarea and images

Table 3: Thematic analysis results of the extracted coupling transition cues.

Category	Code	Definition
Content-level	Extra information is required (**EI**, $N = 19$)	Participants can not proceed with existing information because they do not understand certain knowledge or hope to proceed quickly. Extra information (text and image) is necessary.
	Concerned information appears (**CI**, $N = 18$)	AI mentioned information that the participants were concerned about, be they reactive or proactive, satisfied or unsatisfied.
	Previous goal has (not) been achieved (**GA**, $N = 18$)	Participants have a previous goal, like obtaining extra information, sketching, organizing information, etc., which has (not) been completed or achieved.
	Task step switch (**SS**, $N = 16$)	Participants switch from one predefined step to another.
Behavior-level	Operation in the workspace increases (**OI**, $N = 19$)	Participants initiate or increase their operations in the workspace, like sketching, marking, typing, etc.
	Operate without adding information (**WI**, $N = 10$)	Participants may instigate some operations in the workspace, like scrolling the canvas or conversation histories back and forth, resizing textarea, formatting texts, etc., without contributing solid progress.
	Cool down (**CD**, $N = 3$)	Participants' frequency of communication and operation decrease.

without changing their information. This can imply participants fill the time because they are unable to proceed without AI output.

Cool down (CD ■): Three transition cues were categorized into **CD**, including elongated communication interval, ignoring AI's feedback, and a period without any operation. These cues were also unpredictable and hard to identify because they often take up a certain timescale. We only identified them in *DISC-U* → *SCOL* and *SCOL* → *INDW*, where the coupling of human and AI became looser implicitly over time.

5 Discussion and Future Work

5.1 Guiding the Design of Adaptive AI Behaviors with Coupling Styles

The identified four coupling styles and the dimensions we used to describe them can serve as the basis for designing adaptive AI behaviors. We envision the following two methods for future study to realize adaptive AI behavior in human-AI collaboration:

Behavior-level Adaption: The two dimensions, focused problem and engaged ways, can be translated to certain parameters when designing AI behaviors. For example, "focused problem" can be linked to the range of contexts AI uses for generation. When tightly coupled, AI should analyze collaborator's current focus to stay relevant and avoid adding too much weight to less related information. Contrarily, AI can incorporate a wider range of context from memories to consider more comprehensively or divergently if the focused problems between human and AI are just related in general. Moreover, the dimension "Engaged Way" can impact the output style, where shorter sentences with oral style may be optimal when communicating frequently, while longer paragraphs neatly organized in bullet lists may be preferred when richer information is expected in looser coupling.

Pattern-level Adaption: While behavior-level adaption can be too finely-grained and might be hard to decide the performance of each behavior in a short time, we can also consider combining several behaviors to form behavioral patterns for AI, and adapt in a more coarse way. The setting of behaviors in each pattern can refer to Table 2. Designing behavioral patterns benefits from fewer possible choices when making decisions to transition, and each pattern can be matched with a set of situations to further speed up decision-making. The drawback of this method is also apparent: less flexible and more demanding when fine-tuning the optimal parameters for each pattern.

However, AI does not necessarily need to abide by behaviors defined in current coupling styles originating from human-human collaboration. Our observations suggested that AI can assist in certain tasks that were not described in previous studies. For example, we found participants frequently copying and pasting information when actively discussing with AI (*DISC*), then taking time to organize it. Since current AI models are good at refining information, AI can help refine and organize information while still being able to actively discuss with people, which can not be achieved by a single person in human-human collaboration. Another case is that participants experience a long time waiting without proceeding the design process in *OSTB*, caused jointly by AI's slow generation speed and participants' lack of specific knowledge. Future design of AI behaviors in *OSTB* should seek to optimize user experience in this process by, for example, optimizing AI's workflow to allow pre-generation, apply word-by-word generation, etc.

5.2 Challenges in Transitioning Between Coupling Styles

Although we have identified seven transition cues, enabling AI to autonomously adapt to different coupling styles in collaboration still poses challenges due to the difficulties in recognizing transition cues and determining the timing to initiate transitions.

Among the seven types of transition cues we extracted, **EI**, **GA**, and **CD** are the hardest to recognize because they are 1) subjective, which calls for personalization, 2) commonly seen, which can be hard for AI to distinguish between normal actions and those that foreshadow transitions, and 3) implicit, which can take intensive time to recognize the need for transition and bring heavy delays. Considering the small sample size in this work, future studies can

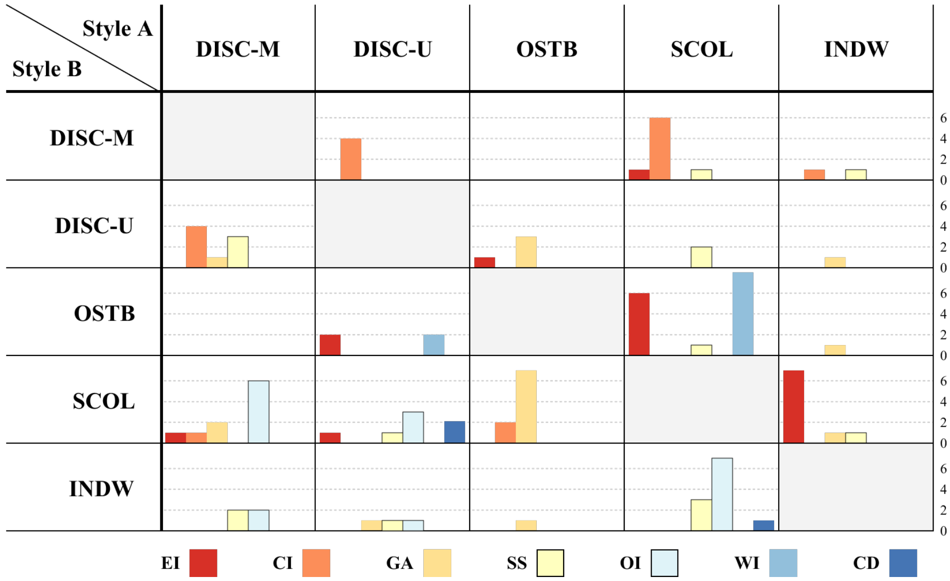

Figure 3: The frequency distribution of all transition cues in 20 types of coupling transition. Style A and Style B at the table head represent the starting and ending coupling style in a transition, i.e., transitioning from Style A (row) to Style B (column).

expand the sample size to generate more data and calculate the percentage of frequencies those commonly seen cues lead to coupling transitions, and feed the expanded and curated data to AI models if prediction is needed in future adaptive AI systems. Besides, although only one out of the seven types is initiated by AI, implying AI catering to human behaviors is a better solution in most cases, we still suggest preserving AI's ability to "interfere" because such interference can lead to the transition to tighter collaboration, supported by previous studies [30]. Also, analyzing human behaviors in real time would inevitably cause delays constrained by current technologies. Future studies should examine the impact of such delays (e.g., [17]) and the threshold of acceptable length of delays.

5.3 Limitations and Future Work

Some limitations in this work ask for future work to strengthen our contributions. First, the data we analyzed is from a lab experiment, which can have discrepancies in human-AI collaboration performances in real-world environments. Conducting user studies with operable prototypes in real human-AI collaborative situations should be considered to ensure the results' practicality. Also, the representative of our participants is limited due to the age range, future studies involving more diverse participants should be considered to eliminate the potential bias brought by age.

Apart from the limitations, our results can potentially be generalized to other scenarios in future work apart from the design process in our experiments. In human-human collaboration, similar coupling styles were observed in collaborative tasks like problem-solving [14, 30], interior design tasks [31], travel planning [1], sense making and information visualization [23], etc., with definitions change accordingly with tasks and scenarios, but key features manifested by the dimensions we concluded remains stable. These studies suggest that similar coupling styles may also exist in different collaborative domains in human-AI collaboration. However, it is still worth noticing that our results are not one-size-fits-all, and could possibly be inapplicable for collaborative tasks that do not exhibit the mixed-focus features.

6 Conclusion

In this work, we propose to view human-AI collaboration as mixed-focus collaboration to inspire future design of adaptive AI. Contextualized in human-AI collaborative design, we analyzed 12 video recordings from six participants, which generated four coupling styles and 7 types of coupling transition cues. Last, we envisioned how these results could inspire future design of adaptive AI, including guiding the design of AI behaviors and how AI can transition between different coupling styles. We hope this preliminary research can offer new insights for the research on adaptive AI.

Acknowledgments

This work was supported by the National Key R&D Program of China (2022YFB3303301).

References

[1] Frederik Brudy, Joshua Kevin Budiman, Steven Houben, and Nicolai Marquardt. 2018. Investigating the Role of an Overview Device in Multi-Device Collaboration. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3173874

[2] Zhuoyi Cheng, Pei Chen, Wenzheng Song, Hongbo Zhang, Zhuoshu Li, and Lingyun Sun. 2025. An Exploratory Study on How AI Awareness Impacts Human-AI Design Collaboration. In *Proceedings of the 30th International Conference on Intelligent User Interfaces* (Cagliari, Italy) *(IUI '25)*. Association for Computing Machinery, New York, NY, USA, 1–16. https://doi.org/10.1145/3708359.3712162

[3] Herbert H. Clark and Deanna Wilkes-Gibbs. 1986. Referring as a collaborative process. *Cognition* 22, 1 (1986), 1–39.

[4] Cristina Conati. 2024. AI-Driven Personalization to Support Human-AI Collaboration. In *Companion Proceedings of the 16th ACM SIGCHI Symposium on*

Engineering Interactive Computing Systems (Cagliari, Italy) *(EICS '24 Companion)*. Association for Computing Machinery, New York, NY, USA, 5–6. https://doi.org/10.1145/3660515.3661324

[5] Paul Dourish and Victoria Bellotti. 1992. Awareness and coordination in shared workspaces. In *Proceedings of the 1992 ACM Conference on Computer-Supported Cooperative Work* (Toronto, Ontario, Canada) *(CSCW '92)*. Association for Computing Machinery, New York, NY, USA, 107–114.

[6] Vanessa Echeverria, Kenneth Holstein, Jennifer Huang, Jonathan Sewall, Nikol Rummel, and Vincent Aleven. 2020. Exploring Human-AI Control Over Dynamic Transitions Between Individual and Collaborative Learning. In *Addressing Global Challenges and Quality Education*. Springer International Publishing, Cham.

[7] Daniel Immanuel Fink, Johannes Zagermann, Harald Reiterer, and Hans-Christian Jetter. 2022. Re-locations: Augmenting Personal and Shared Workspaces to Support Remote Collaboration in Incongruent Spaces. *Proc. ACM Hum.-Comput. Interact.* 6, ISS, Article 556 (Nov. 2022), 30 pages. https://doi.org/10.1145/3567709

[8] Rohit Girdhar, Alaaeldin El-Nouby, Zhuang Liu, Mannat Singh, Kalyan Vasudev Alwala, Armand Joulin, and Ishan Misra. 2023. ImageBind: One Embedding Space To Bind Them All. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (Vancouver, BC, Canada). IEEE, 15180–15190.

[9] Jens Emil Sloth Grønbæk, Juan Sánchez Esquivel, Germán Leiva, Eduardo Velloso, Hans Gellersen, and Ken Pfeuffer. 2024. Blended Whiteboard: Physicality and Reconfigurability in Remote Mixed Reality Collaboration. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 798, 16 pages. https://doi.org/10.1145/3613904.3642293

[10] Carl Gutwin and Saul Greenberg. 1998. Design for Individuals, Design for Groups: Tradeoffs Between Power and Workspace Awareness. In *Proceedings of the 1998 ACM Conference on Computer Supported Cooperative Work* (Seattle, Washington, USA) *(CSCW '98)*. Association for Computing Machinery, New York, NY, USA, 207–216. https://doi.org/10.1145/289444.289495

[11] Carl Gutwin and Saul Greenberg. 2002. A Descriptive Framework of Workspace Awareness for Real-Time Groupware. *Computer Supported Cooperative Work (CSCW)* 11 (2002), 411–446. Issue 3. https://doi.org/10.1023/A:1021271517844

[12] 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* (Newcastle upon Tyne, United Kingdom) *(CHIWORK '24)*. Association for Computing Machinery, New York, NY, USA, Article 9, 14 pages. https://doi.org/10.1145/3663384.3663398

[13] Erzhen Hu, Md Aashikur Rahman Azim, and Seongkook Heo. 2022. FluidMeet: Enabling Frictionless Transitions Between In-Group, Between-Group, and Private Conversations During Virtual Breakout Meetings. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 511, 17 pages. https://doi.org/10.1145/3491102.3517558

[14] Petra Isenberg, Danyel Fisher, Sharoda A. Paul, Meredith Ringel Morris, Kori Inkpen, and Mary Czerwinski. 2012. Co-Located Collaborative Visual Analytics around a Tabletop Display. *IEEE Transactions on Visualization and Computer Graphics* 18, 5 (2012), 689–702. https://doi.org/10.1109/TVCG.2011.287

[15] Hans-Christian Jetter, Jens Gerken, Michael Zöllner, Harald Reiterer, and Natasa Milic-Frayling. 2011. Materializing the query with facet-streams: a hybrid surface for collaborative search on tabletops. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 3013–3022. https://doi.org/10.1145/1978942.1979390

[16] Rutuja Joshi and Klaus Bengler. 2024. Crafting Human-AI Interaction: A Rhetorical Approach to Adaptive Interaction in Conversational Agents. In *Proceedings of the 12th International Conference on Human-Agent Interaction* (Swansea, United Kingdom) *(HAI '24)*. Association for Computing Machinery, New York, NY, USA, 314–322. https://doi.org/10.1145/3687272.3688297

[17] Emily Kuang, Minghao Li, Mingming Fan, and Kristen Shinohara. 2024. Enhancing UX Evaluation Through Collaboration with Conversational AI Assistants: Effects of Proactive Dialogue and Timing. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 3, 16 pages. https://doi.org/10.1145/3613904.3642168

[18] Tianjian Liu, Hongzheng Zhao, Yuheng Liu, Xingbo Wang, and Zhenhui Peng. 2024. ComPeer: A Generative Conversational Agent for Proactive Peer Support. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology* (Pittsburgh, PA, USA) *(UIST '24)*. Association for Computing Machinery, New York, NY, USA, Article 117, 22 pages. https://doi.org/10.1145/3654777.3676430

[19] Inês Lobo, Janin Koch, Jennifer Renoux, Inês Batina, and Rui Prada. 2024. When Should I Lead or Follow: Understanding Initiative Levels in Human-AI Collaborative Gameplay. In *Proceedings of the 2024 ACM Designing Interactive Systems Conference* (Copenhagen, Denmark) *(DIS '24)*. Association for Computing Machinery, New York, NY, USA, 2037–2056. https://doi.org/10.1145/3643834.3661583

[20] Yaxi Lu, Shenzhi Yang, Cheng Qian, Guirong Chen, Qinyu Luo, Yesai Wu, Huadong Wang, Xin Cong, Zhong Zhang, Yankai Lin, Weiwen Liu, Yasheng Wang, Zhiyuan Liu, Fangming Liu, and Maosong Sun. 2024. Proactive Agent: Shifting LLM Agents from Reactive Responses to Active Assistance. arXiv:2410.12361 [cs.AI] https://arxiv.org/abs/2410.12361

[21] Matthew K Miller, Frederik Brudy, Tovi Grossman, George W. Fitzmaurice, and Fraser Anderson. 2023. Peek-At-You: An Awareness, Navigation, and View Sharing System for Remote Collaborative Content Creation. In *Graphics Interface 2023 - second deadline*. https://openreview.net/forum?id=U8p66V2PeEa

[22] Michael Muller, Stephanie Houde, Gabriel Gonzalez, Kristina Brimijoin, Steven I Ross, Dario Andres Silva Moran, and Justin D Weisz. 2024. Group Brainstorming with an AI Agent: Creating and Selecting Ideas. In *International Conference on Computational Creativity* (Jönköping, Sweden). 10 pages.

[23] Thomas Neumayr, Hans-Christian Jetter, Mirjam Augstein, Judith Friedl, and Thomas Luger. 2018. Domino: A Descriptive Framework for Hybrid Collaboration and Coupling Styles in Partially Distributed Teams. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 128 (Nov. 2018), 24 pages. https://doi.org/10.1145/3274397

[24] David Saffo, Andrea Batch, Cody Dunne, and Niklas Elmqvist. 2023. Through Their Eyes and In Their Shoes: Providing Group Awareness During Collaboration Across Virtual Reality and Desktop Platforms. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 383, 15 pages. https://doi.org/10.1145/3544548.3581093

[25] Vildan Salikutluk, Janik Schöpper, Franziska Herbert, Katrin Scheuermann, Eric Frodl, Dirk Balfanz, Frank Jäkel, and Dorothea Koert. 2024. An Evaluation of Situational Autonomy for Human-AI Collaboration in a Shared Workspace Setting. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 300, 17 pages. https://doi.org/10.1145/3613904.3642564

[26] Tony Salvador, Jean Scholtz, and James Larson. 1996. The Denver model for groupware design. *SIGCHI Bull.* 28, 1 (Jan. 1996), 52–58.

[27] Eike Schneiders, Christopher Fourie, Stanley Celestin, Julie Shah, and Malte Jung. 2024. Understanding Entrainment in Human Groups: Optimising Human-Robot Collaboration from Lessons Learned during Human-Human Collaboration. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 152, 13 pages. https://doi.org/10.1145/3613904.3642427

[28] Sabine Sonnentag and Michael Frese. 2012. Dynamic Performance. In *The Oxford Handbook of Organizational Psychology, Volume 1*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199928309.013.0017

[29] Mark Stefik, Gregg Foster, Daniel G. Bobrow, Kenneth Kahn, Stan Lanning, and Lucy Suchman. 1987. Beyond the chalkboard: computer support for collaboration and problem solving in meetings. *Commun. ACM* 30, 1 (Jan. 1987), 32–47.

[30] Anthony Tang, Melanie Tory, Barry Po, Petra Neumann, and Sheelagh Carpendale. 2006. Collaborative coupling over tabletop displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Montréal, Québec, Canada) *(CHI '06)*. Association for Computing Machinery, New York, NY, USA, 1181–1190. https://doi.org/10.1145/1124772.1124950

[31] Philip Tuddenham and Peter Robinson. 2009. Territorial coordination and workspace awareness in remote tabletop collaboration. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Boston, MA, USA) *(CHI '09)*. Association for Computing Machinery, New York, NY, USA, 2139–2148. https://doi.org/10.1145/1518701.1519026

[32] Xinyu Wang, Bohan Zhuang, and Qi Wu. 2024. ModaVerse: Efficiently Transforming Modalities with LLMs. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (Seattle, WA, USA). IEEE, 26606–26616.

[33] Kexin Bella Yang, Vanessa Echeverria, Zijing Lu, Hongyu Mao, Kenneth Holstein, Nikol Rummel, and Vincent Aleven. 2023. Pair-Up: Prototyping Human-AI Co-orchestration of Dynamic Transitions between Individual and Collaborative Learning in the Classroom. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 453, 17 pages. https://doi.org/10.1145/3544548.3581398

[34] Michelle Zhao, Reid Simmons, and Henny Admoni. 2022. The Role of Adaptation in Collective Human–AI Teaming. *Topics in Cognitive Science* (2022), 1–33. https://doi.org/10.1111/tops.12633

[35] Jiayi Zhou, Renzhong Li, Junxiu Tang, Tan Tang, Haotian Li, Weiwei Cui, and Yingcai Wu. 2024. Understanding Nonlinear Collaboration between Human and AI Agents: A Co-design Framework for Creative Design. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 170, 16 pages. https://doi.org/10.1145/3613904.3642812

A The Descriptive Coupling Style Coding Scheme

Table 4: The descriptive coupling styles coding scheme adapted from Neumayr et al. [23]. DP and D were excluded because they would not appear in our experiment.

	Focused Problem	Engaged in Communication	Engaged in Operation	Definition
DISC	Same	Very frequent	Sometimes	Active **discussion** between two collaborators about the task. Limited system operation (e.g., scrolling the canvas)
VE	Same	Frequent	Frequent/Very infrequent	**View engaged.** One collaborator is actively working in the workspace, another watches and engages in conversation and commenting on the observed activities, but is not operating with the system
SV	Similar	Very infrequent	Infrequent	**Sharing of the same view** of an item. Collaborators look at the same display that shows the information
SIDC	Similar	Sometimes	Frequent	**Sharing of the same information but using different displays for coordinated exploration**. Collaborators view the same information item but use different displays for coordinated exploration
SIDD	Similar	Very infrequent	Very frequent	**Sharing of the same information but using different displays**. Collaborators view the same information item but use different devices and are not engaged in active conversation.
SSP	Related	Infrequent	Very frequent	Work is shared to solve the **same specific problem**. Users read different information items from a shared set.
SGP	Related in general	Very infrequent	Very frequent	Work on the **same general problem** but from different starting points.