Humanities and Social Sciences Communications

Article in Press

https://doi.org/10.1057/s41599-026-06738-7

Large language model tools as catalysts for collective cognition in collaborative new-product development: a quasi-experimental study

Received: 16 January 2025

Accepted: 10 February 2026

Cite this article as: Zhang, D., Luo, S., Liu, Y. *et al.* Large language model tools as catalysts for collective cognition in collaborative new-product development: a quasi-experimental study. *Humanit Soc Sci Commun* (2026). https://doi.org/10.1057/s41599-026-06738-7

Deyin Zhang, Shijian Luo, Yue Liu, Xiyuan Zhang, Yuqi Hu, Jianan Shen, Peiqi Yi, Kayu Zhao & Wei Liu

Large language model tools as catalysts for collective cognition in collaborative new-product development: a quasi-experimental study

ABSTRACT

Modern organisations frequently face complex cognitive challenges in Collaborative New-Product Development (Co-NPD), particularly when integrating dispersed knowledge and coordinating work across different development phases. To investigate how Large Language Models (LLMs) influence collective cognition and collaborative processes, this study introduces the 2I2A model, which defines four collaboration spaces and eight associated communication dimensions. A mixed-methods design was adopted, combining quantitative analysis of collaborative behaviours with semi-structured interviews analysed using grounded theory. The findings indicate that LLMs mainly support the early stages of Co-NPD by expanding collective cognitive boundaries, improving knowledge integration and facilitating idea generation. However, their contribution to deeper analytical reasoning, negotiation and solution integration is more limited. The grounded theory analysis additionally highlights potential drawbacks, including reduced collaborative naturalness and a tendency toward over-reliance on LLM-generated suggestions. Overall, the study suggests that the 2I2A model offers a useful framework for examining how collective cognition develops in Co-NPD and clarifies both the potential and the boundaries of LLM assistance in collaborative innovation.

Keywords

Co-NPD, LLMs, cognition model, collective cognition

1 Introduction

Collaborative new product development (Co-NPD) is a systematic innovation approach that encourages the formation of a shared cognitive framework within and across organizational teams (Beverland et al., 2016) . By establishing this common foundation, Co-NPD improves collaboration efficiency and enhances overall creativity (Marzi & Balzano, 2025)(Christiansen & Varnes, 2009). Co-NPD emphasises the transformation of individual cognition into collective cognition among team collaborators (Lobo et al., 2025). This transformation expands cognitive boundaries and strengthens teams' shared understanding of development goals and design requirements (Kozar, 2010). Through this cognitive convergence, collaborators can integrate diverse knowledge resources and generate more novel ideas and solutions (Järvelä et al., 2016). In practice, collective cognition functions as a dynamic system that supports knowledge sharing and idea dispersion. This system facilitates team decision-making and improves the quality of collaborative outcomes (London & Singh, 2013; Razmerita et al., 2014).

Recent developments in Artificial Intelligence (AI) and the rapid popularity of Large Language Models (LLMs) tools have brought vast potential for Co-NPD. Given their strong natural language processing and reasoning capabilities, LLMs can assist teams in recognising, defining and addressing complex innovation problems (Ge et al., 2025; Zamfirescu-Pereira et al., 2025). Empirical research on LLM-assisted team tasks suggests that these tools can enhance idea generation, coordination, and information structuring in complex collaborative design scenarios (Qiu & Jin, 2025; Suh et al., 2024; Feng et al., 2025). Current studies have preliminarily explored the benefits of utilising LLMs at the individual level in education, technology and business domains (Omidvar Tehrani et al., 2024; Nejjar et al., 2023; Murtaza et al., 2024). However, despite these insights, there remains a need for systematic and fine-grained examination of how LLMs influence collective cognitive flow in complex Co-NPD scenarios. In particular, the practical effects and applicability of LLMs in dynamically adapting to and analysing complex innovation tasks are not yet well understood (Woolley, 2025).

This study investigates the potential of LLMs in Co-NPD, focusing on how these tools support the construction of collective cognition and influence collaboration quality and efficiency. Through empirical analysis, we aim to uncover the mechanisms by which LLMs shape cognitive and collaborative processes in complex innovation settings. This study addresses the following research questions:

RQ1: How do LLMs influence the construction of team collective cognition during Co-NPD?

RQ2: How do LLMs alter team members' collaboration patterns during Co-NPD?

RQ3: What are the practical values and limitations of LLMs during Co-NPD?

To answer the above research questions, the main contributions of this paper are as follows:

1. It proposes and validates the 2I2A model for analysing cognitive processes in Co-NPD, offering a structured perspective for examining collective cognition.
2. It advances understanding of the role of LLMs in collaborative processes and collective cognitive flow.
3. It identifies the practical value and boundaries of LLMs in Co-NPD, particularly their dual role in supporting innovation and enhancing collaboration efficiency.

The remainder of this paper is organised as follows. Section 2 reviews the theoretical background, including collective cognition in collaborative contexts (2.1), Co-NPD and the cognitive interaction model (2.2), the evolution of intelligent technologies for collaborative work (2.3), and the role of LLMs in collaborative innovation (2.4). Section 3 outlines the research methodology, including participant recruitment (3.1), experimental design and tasks (3.2), and the quantitative and qualitative analytical procedures (3.4 and 3.5). Section 4 presents the results, including both statistical analyses and grounded-theory findings. Section 5 discusses the implications of the findings, relating them to existing research and identifying practical applications and limitations. Finally, Section 6 concludes the study, summarising the key contributions and outlining directions for future research.

2 Literature review

2.1 Collective cognition in collaborative contexts

Team collaboration is an essential driving force for enterprise innovation (Cohn et al., 2024). Collaboration involves multiple individuals continuously defining problems, constructing ideas and providing feedback. This iterative process reflects ongoing interaction between collaborators' individual cognitions and the team's collective cognitive system (Kozar, 2010). As an important indicator of effective collaboration (Järvelä et al., 2016), collective cognition is commonly described as a dynamic system that supports the construction, coordination and transformation of shared understanding (Oygür, 2018). Collective cognition has two core characteristics: First, it emerges through bottom-up interactions among individuals' cognitive contributions (DeChurch & Mesmer-Magnus, 2010). Second, it is situated within shared interpretive frames developed through interaction, rather than within any single collaborator's mind (Danish et al., 2020;

Gibson, 2001; Järvelä et al., 2016). From this perspective, collaboration becomes a constructivist and reflective group practice (Yang et al., 2024). Shteynberg (2023) suggests that shared cognitive representations can amplify collective attention, strengthen relational ties and increase collaborative engagement.

The current academic research on collective cognition mainly covers the following aspects: First, Shared Intention and Cognitive Evolution. Asarnow's research (2020) reveals that shared behaviour does not require shared intentions, and the taste for sociality and cognitive-conative theory of mind can facilitate shared agency. Heintz & Scott-Phillips's research (2023) show that diverse human expressive behaviours arise from interrelated cognitive abilities. Their work further suggests that different expressive acts function as sub-functions of a unified capacity for ostensive communication. Second, Collective Beliefs and Socially Evaluative. Vlasceanu et al (2024) argue that beliefs can be understood as a multidimensional system of mental representations that spans three cognitive structures at both individual and collective levels. They further emphasise that such beliefs tend to be highly resilient to interventions. Woo's research (2023) shows that the human capacity for mentalizing is enhanced in social evaluation contexts, especially when there is a need to assess the behaviour of others as potential social partners. Third, Cognitive Adaptation and Interaction Dynamics. Iskra et al (2024) observes that sports decision-making research has relied heavily on static and simulated options but lacks cognitive-action interaction dynamics. Galesic's research (2023) highlights that collective adaptation transcends the notion of collective intelligence. Several key aspects of real-world collective behaviour include path dependence, collective transience, and aimless exploration of alternative worlds.

Additionally, Complexity and Cognitive Overload. Chen et al. (2023) identify that task complexity can be measured by element interactivity, which integrates information structure and knowledge in learners' long-term memory. Patel & Alismail's research (2024) points out that emotional and motivational factors have a significant influence on cognitive load. Their study also shows that regulating negative emotions can improve learning outcomes. Lastly, Cognitive Analysis of Technological Interventions. Wang et al. (2023) emphasize that mobile technology in primary and secondary education positively impacts students' cognitive, affective, and behavioural learning outcomes. Hao's research (2024) shows that generative artificial intelligence can reduce cognitive burden in complex situations. However, the study also notes several challenges, including technology dependence, inherent biases and limited situational creativity.

However, most existing studies focus on traditional forms of team collaboration and mainly examine how cognitive mechanisms vary across specific contexts (Cooke et al., 2024; Carraro et al., 2025). They offer limited insight into the cognitive interactions that unfold in complex, dynamic and multi-stage collaborative scenarios. In addition, although new AI technologies and LLM tools such as ChatGPT are now widely used, their influence on team collective cognition has received relatively little attention (Li et al., 2024; Gao et al., 2024). In particular, little is known about how LLMs support knowledge integration or help teams adapt to complex innovation tasks (Woolley, 2025).

2.2 Co-NPD and cognitive interaction model

Co-NPD is a structured innovation model grounded in cognitive interaction and driven by team-based knowledge integration (Beverland et al., 2016; Christiansen & Varnes, 2009; Zahra et al., 2007). It emphasises the complex, iterative work that occurs from concept generation to final product realisation. This process depends on continuous interaction and cognitive integration among team members. In contrast to traditional linear product development, Co-NPD highlights dynamic knowledge sharing, timely feedback and the joint use of collective expertise within and across teams (Beverland et al., 2016). Corallo et al. (2012) point out that this collaborative model is particularly suitable for innovation under conditions of high uncertainty and complexity. It enables teams to overcome limitations in individual knowledge, experience and thinking by facilitating effective cognitive interaction. These mechanisms help reduce cognitive inertia and design fixation by expanding the organisation's knowledge base (X. Jin et al., 2024). A broader knowledge base, in turn, supports the development of more feasible and innovative solutions (Nagaraj et al., 2020).

Cognitive interaction is a key capability that shapes innovation in Co-NPD (Dell'Era et al., 2020). Its core involves transforming independently held knowledge into shared understanding through knowledge fusion and collaborative communication (Razmerita et al., 2014). This process expands cognitive boundaries and helps team members form a more comprehensive understanding of development goals and innovation requirements. Kyriakopoulos & De Ruyter's research (2004) shows that success in Co-NPD depends on establishing a shared cognitive framework within the team. Such a framework not only improves collaboration efficiency but also supports greater creativity.

However, cognitive design processes are often difficult to observe or formalise. They are frequently intuitive, serendipitous and partly subconscious (J. H. Lee et al., 2020). To study these processes in greater depth, researchers have proposed several cognitive models. The Design Team Cognition (DTC) model conceptualises collaborative knowledge as a type of transactive memory and emphasises cognitive, collaborative and creative processes (J. H. Lee et al., 2020). It comprises two knowledge bases: emergence and sharedness, which relate to communication and performance; and distributed knowledge, which relates to creativity and cognition.

Badke-Schaub et al. (2007) developed another model that identifies five key elements involved in individual cognitive transitions: task, process, team, competence and environment. Their model also highlights three factors that influence the quality of mental models: sharedness, accuracy and importance. Building on Sternberg's research (1998), Stempfle & Badke-Schaub (2002) proposed a cognitive operation model that defines four fundamental operations involved in problem-solving: generation, exploration, comparison and selection. These operations form part of the cognitive interaction process, offering a clear framework for expanding and narrowing problem spaces. Ostergaard & Summers (2008) developed an analytical framework for collaborative design that identifies six key factors: team composition, communication, distribution, design approach, information and problem type.

Gibson (2001) provided an early analysis of how knowledge accumulates within teams and developed a widely cited model of collective cognition. The model outlines four interchanging stages of collaboration: accumulation, interaction, examination and accommodation. It also identifies feedback, consensus and conflict and other collaborative processes as key processes that shape cognitive flow within groups. These studies provide an important theoretical foundation for team cognitive interaction in Co-NPD.

Gibson's (2001) framework provides a foundational understanding of the collective cognition cycle through four interchanging "stages". However, it faces challenges when applied to highly complex and dynamic tasks such as Co-NPD. For example, the "Accumulation" stage offers a broad description of knowledge gathering. However, it provides less guidance on proactive, goal-oriented identification of problems and information, which is important for innovation. The “Examination” stage brings together general discussion and more structured analytical work. This overlap may make it harder to distinguish systematic, evidence-based reasoning, particularly in contexts where AI-assisted analysis is involved.

To address these gaps, we introduce the 2I2A framework. It refines Gibson’s stages into four concurrent cognitive spaces: Identification, Interaction, Analysis and Accommodation. This structure enables more fine-grained observation by linking team behaviours to eight collaborative dimensions: Information Perception, Information Selection and Accumulation, Structuring of Questions and Ideas, Cognitive Exchange, Evaluation, Negotiation and Interpretation, Solution Accommodation, and Creative Output. These dimensions allow us to examine how LLMs contribute to different aspects of cognitive and collaborative activity without overstating their role.

2.3 Evolution of intelligent technologies for collaborative work

Intelligent technologies that support collaborative work have evolved along several technological pathways. Some systems focus on structured Knowledge Management (Riyadh et al., 2021; Cha et al., 2015). Others enhance analytical capacity through data-driven methods (Sarker, 2021; Alaskar et al., 2024). More recent approaches incorporate interaction designs that enable richer and more flexible human-machine communication (Z. Zhang et al., 2025; Xu et al., 2025). Examining these developments provides the broader context for understanding why large language models have become increasingly relevant to collaborative intelligence.

Early intelligent systems, including Knowledge Management Systems (KMS) (Alavi & Leidner, 2001), expert systems (Shu-Hsien Liao, 2005), and Computer-supported Cooperative Work (CSCW) systems (Pratt et al., 2004), relied primarily on structured knowledge bases and predefined rules. These tools supported tasks such as solution evaluation and risk assessment through rule-based outputs and formalised knowledge structures (Cowan, 2001; Fan et al., 2008; Edwards et al., 2005). While KMS and expert systems facilitated knowledge reuse and reduced cognitive load associated with information retrieval, their capacity was limited. They struggled to process unstructured language or tacit knowledge generated during collaboration (Tseng, 2008). CSCW systems were more effective at improving communication and information sharing, but contributed less to deeper forms of cognitive alignment (Gross, 2013). As a result, the influence of these early systems on collaborative thinking remained constrained by their dependence on structured information.

With the expansion of data resources and the advancement of computational methods, organisations increasingly adopted data-driven tools to enhance analytical capability. Business Intelligence (BI) systems support historical reporting, trend identification, and performance evaluation across sectors (Cody et al., 2002). They have been widely applied in sectors such as healthcare and education to support trend analysis, performance evaluation, and data-driven decision-making (Alkhwaldi, 2024; Alkhwaldi et al., 2025). This analytical foundation plays a role in organisational learning and strategic consensus-building. Complementing BI, more sophisticated algorithmic methods, such as Machine Learning (ML) (Jordan & Mitchell, 2015) and predictive analytics (Shmueli & Koppius, 2011) further enhance strategic foresight. ML identifies complex patterns and provides predictive indicators for decision-making (Trovato et al., 2025), while predictive analytics forecasts future trends and supports scenario-based planning (Ugbebor et al., 2024). These algorithmic tools effectively expand the analytical space available to teams. However, these tools are primarily oriented toward quantitative interpretation and insight delivery, focusing on what the data indicates (Herm et al., 2023). Consequently, they offer limited direct support for real-time knowledge construction, negotiation, or co-creation within teams (Nahar et al., 2022). Their analytical logic operates on a different dimension from the generative and linguistic capabilities introduced by LLMs.

To better understand the cognitive mechanisms underlying these technological interventions, research in human-computer interaction (HCI) offers a theoretical bridge between earlier structured systems and contemporary AI (Carroll, 1997; Liu, 2024; Ho & Vuong, 2025). This body of work examines how users interact with technological systems and how interface design shapes cognitive processes (Hurtienne, 2009; Chalmers, 2003; Zhang & Soergel, 2014). Studies show that interface transparency, levels of automation and feedback design affect attention allocation, verification behaviours and reliance on system outputs (Parasuraman et al., 2000; J. D. Lee & See, 2004; Parasuraman & Manzey, 2010; Hoff & Bashir, 2015). Higher levels of automation can increase susceptibility to automation bias (Onnasch et al., 2014). In addition, the format of information presentation influences both cognitive load and judgment strategies (Wickens, 2002; Scharowski et al., 2023). Research on distributed cognition highlights the role of digital representations in enabling teams to externalise and coordinate cognitive processes (Fiore & Wiltshire, 2016; J. H. Lee & Ostwald, 2025). Studies on trust calibration further show that users often misjudge system reliability, resulting in either overreliance on or unnecessary rejection of automated recommendations (J. D. Lee & See, 2004; Lebiere et al., 2021; Carragher et al., 2024). Collectively, these insights establish a comprehensive foundation for understanding how artificial intelligence influences team reasoning processes. They help explain the evolving dynamics of the human-tool relationship when natural language interfaces are introduced (Andrews et al., 2023).

These technologies provide valuable foundations in knowledge management, data analytics, and cognitive support. However, they still rely heavily on structured information and relatively fixed interaction patterns (Salma et al., 2025; Zhou et al., 2024). This reliance limits their ability to address the open-ended and semantically rich exchanges required in Co-NPD. Advances in natural language processing have gradually introduced tools that can engage with unstructured language, generate diverse responses, and support more flexible forms of reasoning during collaborative work (Hirschberg & Manning, 2015; Huang et al., 2025). These developments complement earlier systems and illustrate the ongoing evolution of technological support for collaborative work. In this context, LLMs have attracted increasing attention as a promising tool for supporting open-ended, language-rich collaboration in Co-NPD.

2.4 Large language models in collaborative innovation

Researchers and practitioners see IT infrastructures as important tools for the successful NPD. These infrastructures are essential in facilitating knowledge integration and optimising the innovation process (Y. Chen et al., 2015). As such, they are considered an important source of competitiveness in new product development (Qin et al., 2021). Recently, the Co-NPD innovation approach has been extended to the field of human-computer collaboration. This extension has been facilitated by the development of deep learning and natural language processing technologies, alongside the widespread use of LLM applications. With its excellent natural language understanding and reasoning capabilities, the LLM tools help teams efficiently identify, define and solve complex problems in the new product development process. Research has shown that LLMs are helpful in information acquisition (Vu et al., 2024), copywriting (Wohllebe & Lagodka, 2024), software development (Omidvar Tehrani et al., 2024), data analysis (Nejjar et al., 2023), education and training (Murtaza et al., 2024), and customer service (Kolasani, 2023).

In human-computer collaborative environments, LLMs are believed to have the potential to enhance human cognitive abilities (Demetriadis & Dimitriadis, 2023). Kernan Freire et al. (2024) found that LLMs can help knowledge creation and sharing in the factory production process; Wei et al. (2024) explored that collaboration with LLMs enhanced learners' willingness to learn with emotional satisfaction, as well as their tendency to collaborate and their academic performance; Liu et al. (2024) found that when serving as facilitators of teamwork, LLMs are more helpful in fostering creative thinking in children; Memmert & Tavanapour (2023) found that LLMs can provide significant stimulation of cognitive stimulation, facilitating the generation and iteration of more novel ideas in collaboration; Gonzalez (2024) found that the role of LLMs can be extended to facilitate team brainstorming, rather than being limited to generating ideas. Studies also found that Chat GPT was able to improve mutual trust in collaboration when participating in human-computer collaboration (Ye et al., 2023). These findings suggest that LLMs stimulate cognitive processes and optimise information transfer and knowledge integration, creating more potential value for collaborative innovation.

Although existing studies have revealed many potential applications of LLMs, most still focus on human-computer collaboration and cognitive assessment at the individual level (Xu et al., 2025; Kang et al., 2025). Research on how LLMs influence team-level processes in complex innovation tasks remains limited (Burton et al., 2024). In particular, little is known about whether LLMs can support collective cognition, optimise team collaboration patterns, and enhance the efficiency and creativity of collaborative innovation (Marzi & Balzano, 2025). This paper takes Co-NPD as the research background and investigates the role of LLMs in team collaboration and innovation to fill these research gaps. By analysing their functions and challenges in cognitive interactions, this paper aims to reveal how LLMs empower collaborative innovation teams. It also seeks to provide theoretical references and practical guidance for future research in related fields.

3 Methods

This study employed a sequential mixed-methods design, following Creswell & Clark's (2017) suggestion. The process began with qualitative discourse analysis to encode collaborative data, followed by quantitative methods to examine inter-group differences. Finally, qualitative grounded theory was applied to explore collaborators' perceptions.

By integrating both quantitative measures of team collaboration and semi-structured interviews, this study ensured a comprehensive understanding of the Co-NPD process. The quantitative analysis provided an objective assessment of collaborative behaviours and cognitive patterns across different stages, while the qualitative interviews offered rich insights into participants' subjective experiences. This combined approach captured both observable outcomes and underlying cognitive processes, providing deeper insights than single-method designs could offer.

3.1 Participants

We posted experimental recruitment on several university forums. Eventually, we recruited 44 (17 males, 27 females) senior undergraduate and junior master's students from design and engineering disciplines, with an average age of 21.1 years ($SD = 0.71$). A summary of the detailed demographic characteristics and educational level of the participants is provided in **Table 1**. All participants had prior experience using LLM tools, having interacted with them at least once a week over the three months preceding the experiment. This ensured that participants were familiar with typical LLM interfaces and functionalities. Additionally, participants had more than three years of study in design or three years of experience in product development.

Doctoral students and professional designers were intentionally excluded. This decision was based on the significant experience gap between them and the participants (undergraduate and junior master's students). This difference in prior knowledge and expertise could have confounded the assessment of LLM support in collaborative cognition. All participants performed the tasks under the same time constraints, ensuring comparable levels of time pressure across teams.

Table 1. Participant Demographics

Characteristic	Category / Value	N	Percentage (%)

Total Sample Size	N = 44	44	100%
Gender	Male	17	38.64%
	Female	27	61.36%
Age	Mean Age (SD)	21.1 (0.71)	N/A
Educational Level	Undergraduate Students	39	88.64%
	Junior Master's Students	5	11.36%
Field of Study	Industrial Design	25	56.82%
	Visual Design	17	38.64%
	Environment Design	2	4.55%

This study involved human participants, and all procedures adhered to institutional ethical standards. Before the experiment, all participants signed a written informed consent form and were fully informed of the study's purpose, procedures, potential risks, and their right to withdraw from the study at any time without penalty. No personally identifiable or sensitive information was collected during the experiment or interviews.

3.2 Quasi-experimental design and design task

The 44 participants were divided into 22 collaborative teams of two people per team and further assigned to two experimental conditions (LLM-assisted or non-LLM), with 11 teams per group. All teams were provided with paper, pencils, and erasers, and each collaboration session was strictly limited to 90 minutes. Participants collaborated to design an innovative product for children. They were required to verbalise their reasoning using the think-aloud method (Eccles & Arsal, 2017). This procedure enabled the collection of naturalistic collaborative communication for subsequent protocol analysis (Charters, 2003). All verbal communication was audio-recorded.

The two groups differed only in whether LLM assistance was available during collaboration. Participants in the LLM-assisted condition were instructed to use ChatGPT-4.0 or Kimi (running on a MacBook Pro laptop) during collaboration and were prohibited from accessing other websites or search engines. To ensure transparency and consistency of the intervention, participants were permitted to interact freely with the LLM throughout the task. They could use it for information search, idea expansion, problem clarification, or solution refinement as needed. All prompts were generated by the team members themselves without researcher intervention, and multi-turn dialogue was allowed to replicate naturalistic use. LLMs served solely as text-based assistants, with no external web browsing allowed. Teams were also instructed to refrain from seeking information unrelated to the design task. By contrast, teams in the non-LLM condition were not allowed to use any digital devices during collaboration.

Although participants worked in pairs, all quantitative analyses were performed at the team level rather than the individual level. Each team contributed one aggregated value per variable, and no individual-level measurements were used in any inferential test. Accordingly, the statistical assumptions regarding independence were satisfied at the team level, and the analytical approach aligns with standard practices in team-based cognitive and collaboration research.

3.3 Protocol analysis and 2I2A model

Under experimental settings, researchers see verbal communication as "construct mental spaces, the relationships between them, and the relationships between the elements within them" (Fauconnier, 1994). To explore the construction and interaction process of collective cognition, this paper uses Protocol Analysis to code and analyse data from participants' verbal communication during the experiment. Protocol Analysis is a widely used methodology in applied psychology, cognitive science, and behavioural analysis. It collects verbal exchanges from participants to reveal the process of cognitive construction (Ericsson, 2017).

Data coding builds on the collective cognition model proposed by Gibson (2001), which specifies four iterative collaborative spaces and eight dynamic transition processes in collaboration. Unlike Gibson's sequential "stages", the 2I2A model defines four concurrent "Spaces". This approach allows for simultaneous tracking of collaborative cognition and supports the operationalization of AI-assisted interactions in Co-NPD. Initially, we faced challenges when applying Gibson's framework due to its complexity and the overlapping nature of some definitions. To address these limitations, the prototype was refined and developed into the 2I2A (Identification, Interaction, Analysis, Accommodation) model, as shown in **Figure 1.**

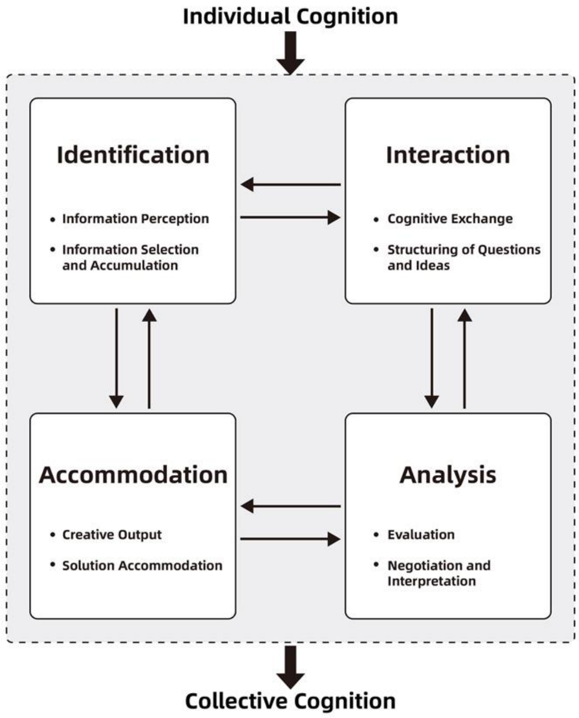

Figure 1. The 2I2A Model of Collective Cognition

The 2I2A model conceptualizes collaborative communication within the notion of different spaces, highlighting the cognitive interactions among collaborators within the team (Razmerita et al., 2014). It specifies four collaborative Spaces (Identification, Interaction, Analysis, and Accommodation) as first-level dimensions. These are further subdivided into eight types of collaborative communication (cf. **Table 2**). In this study, the 2I2A model served as the foundation for coding and quantifying collaborative interactions. Each observed exchange was mapped to one of the four Spaces and categorized into one of the eight sub-dimensions. This operationalization ensured consistency in coding collaborative behaviours across the four Spaces and eight communication types. In this way, the 2I2A model both operationalizes and extends Gibson's framework, offering theoretical clarity and practical utility in assessing multi-dimensional cognitive support.

Table 2. Dimension and Definition of the 2I2A Model

Dimension	Sub-dimension	Definition	Sources
Identification Space	Information Perception	Receive task-related information and external	(M. H. Kim et al., 2007)

		environmental cues.	
	Information Selection and Accumulation	Collect and aggregate relevant information required for the mandate.	(Lloyd & Scott, 1994) (M. H. Kim et al., 2007)
Interaction Space	Structuring of Questions and Ideas	Sort, organise and logically manipulate problems and ideas.	(Chan, 1990) (J. Lee et al., 2014)
	Cognitive Exchange	Deepen understanding through the exchange of information, experiences and feelings.	(Stauffer & Ullman, 1991) (Liikkanen & Perttula, 2009)
Analysis Space	Evaluation	Interpret and evaluate its relevance and implications for decision-making.	(Y. Jin & Benami, 2010) (J. Kim & Ryu, 2014)
	Negotiation and Interpretation	Reconcile differences and explain each other's views to reach agreement.	(J. Kim & Ryu, 2014) (Ball et al., 2004)
Accommodation Space	Solution Accommodation	Integrate diverse inputs to formulate a coherent and viable solution.	(Leblebici-Başar & Altarriba, 2013)
	Creative Output	Generate innovative and actionable results based on synthesized information.	(M. H. Kim et al., 2007) (Stauffer & Ullman, 1991)

Figure 2 provides a concise example of how the 2I2A model was applied in practice. The episode begins with the two participants reviewing GPT's suggestion and working to clarify the form and functional logic of the proposed device. At this stage, their discussion focuses on structuring an initially ambiguous idea. This segment was coded as Interaction Space, as the participants were reorganizing the idea, shaping the problem structure, and refining the design question. These activities reflect Structuring of Questions and Ideas, rather than the acquisition of new information.

As the conversation evolves, the focus shifts from structuring to assessing the practical feasibility of the idea, signaling a move into the Analysis Space. One participant raised a concern about how the device would function when the child was not nearby. This reflects Evaluation, as it involves assessing constraints and examining the practical implications of the proposed design. Building on this insight, the other participant suggested embedding monitoring and reward functions into the child's digital device, marking the transition into the Accommodation Space. At this stage, the previous reasoning was synthesized into an actionable concept. The contribution was coded as Creative Output.

After forming this preliminary idea, the participants revisited GPT's list of reference products and extracted functional features such as app-based operation, gamified reward systems, and interactive elements like virtual pets or screens. This marks the movement into the Identification Space, specifically corresponding to Information Selection and Accumulation. The participants were expanding their information base by gathering and consolidating task-relevant inputs. This differs from the earlier Interaction segment, which focused on reorganizing an existing idea rather than adding new information.

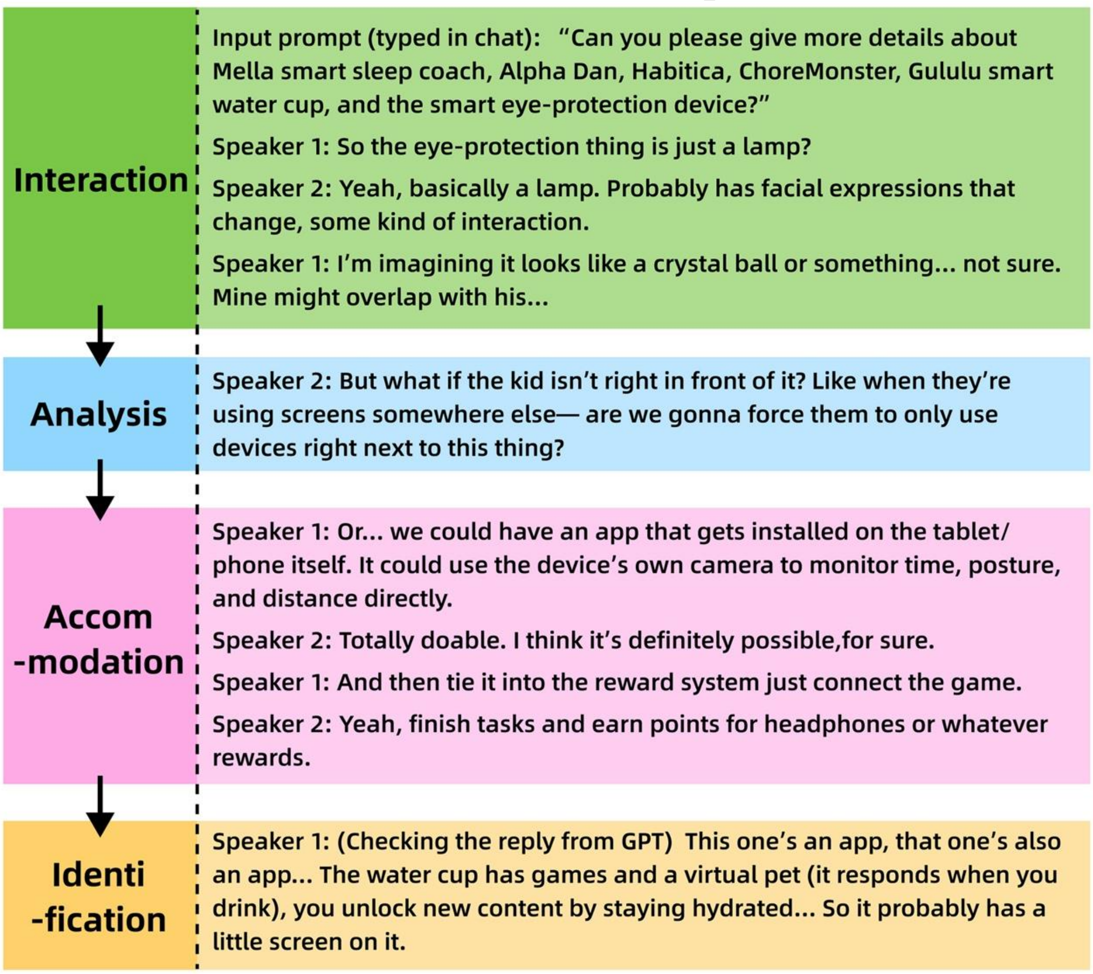

Figure 2. Illustrative Example of 2I2A Coding Across Collaborative Spaces

The development of the 2I2A model aims to achieve three key objectives: first, to clarify and quantify collaborators' interactions within different collaborative spaces; second, to explore how collaborators make cognitive transitions between collaborative spaces and the structure of cognitive interactions; and third, to provide new perspectives for understanding, evaluating, and optimising collaboration and decision-making processes.

To ensure the rigour of the protocol analysis, all collaborative communication (approximately 36 hours across 22 teams) was audio-recorded, transcribed, and independently coded by two coders (the first and second authors) to maintain objectivity. A random 20% subset of the collaborative data was double-coded to assess inter-coder reliability. Cohen's Kappa for the four collaborative Spaces was 0.64 ($p < 0.001$), indicating moderate to substantial agreement. This level of reliability is consistent with widely accepted benchmarks for complex and context-dependent discourse coding (Landis & Koch, 1977; Lombard et al., 2002). The frequency counts for the eight communication sub-dimensions also showed strong correspondence between coders (Spearman's $\rho = 0.81$, $p < 0.001$). All discrepancies were resolved through discussion to reach full consensus before coding the remaining data.

3.4 Qualitative data collection: semi-structured interviews

To gain rich insights into the mechanistic "why" behind the quantitative results, semi-structured interviews were conducted with all 22 participants assigned to the LLMs group (11 teams). These interviews were executed within one week of the main experiment and lasted approximately 35 minutes each. The interview outline focused on three main aspects: attitudes and perceptions towards LLMs, the specific ways in which LLM tools were used, and the perceived role of LLMs during Co-NPD (cf. **Table 3**). During the interviews, the researcher employed probing questions to explore potentially valuable responses in detail.

The qualitative data analysis was subsequently guided by the principles of Grounded Theory, utilising continuous data collection, initial open coding, and axial coding (Corbin & Strauss, 1990). This structured approach ensured the rigorous development of emergent concepts and facilitated the achievement of theoretical saturation. While all 22 interviews were conducted, the systematic analysis process adhered strictly to the continuous comparative method of Grounded Theory. This approach focused on achieving the required depth and rigour in concept development. Data analysis revealed that the core conceptual themes were nearing saturation after the 14th interview transcript. The assessment of thematic saturation followed established criteria widely adopted in qualitative research, whereby saturation is reached when no new conceptual codes emerge from consecutive interviews (Bowen, 2008; Urquhart et al., 2010). The 15th and 16th transcripts yielded no additional codes or relationships, confirming that theoretical saturation had been achieved. Therefore, a final dataset of 16 interview transcripts was deemed sufficient and obtained for the in-depth grounded theory analysis.

Table 3. Interview Outline and Questions

Interview outline	Interview Questions
Attitudes and perceptions towards LLM tools	What are the advantages and disadvantages of LLMs for collaborative new product development?
	Do you rely on these tools when collaborating during Co-NPD?
	What is your attitude towards these tools during use, and has there been any change?
	Which features of LLMs are most easily accepted by the team, and which may be resisted?
How LLM tools were used during Co-NPD	Has the introduction of LLMs changed the way your team communicates and collaborates?
	Have LLMs helped you overcome creative barriers during Co-NPD?
	What impact do LLMs have on your design thinking or problem-solving approach?
	Have LLMs helped you expand your thinking or provided new perspectives?
The role of LLM tools during Co-NPD	What role do LLMs play in different stages of collaboration?
	Do you prefer to view LLMs as a member of the team or as a tool during Co-NPD?
	Will there be any changes in cooperation and conflict when LLMs

	participate?
	Can LLMs help alleviate certain conflicts?
	Has design process or design thinking changed when LLM participated?

Two researchers iteratively analysed the interview data using Nvivo 14 software, following the Grounded Theory (Lane & Seery, 2011; Tan, 2010). Grounded theory is a comprehensive and structured analysis method conducted in three stages (cf. **Figure 3**).

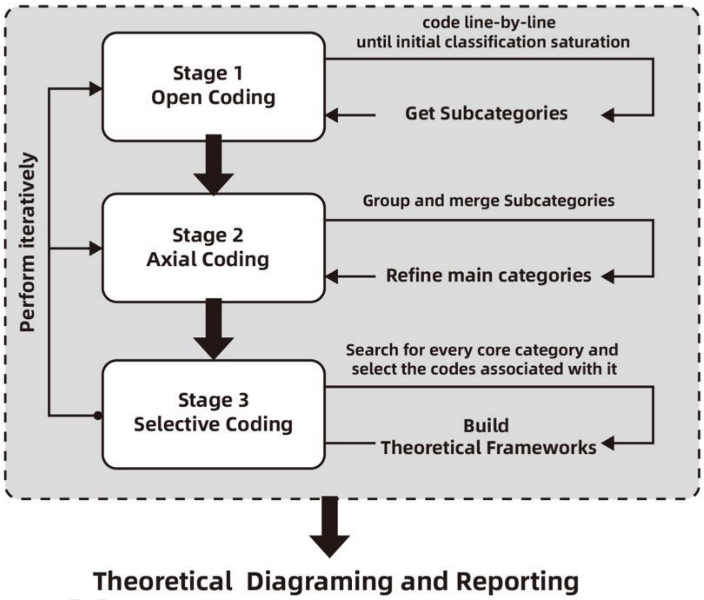

Figure 3. Stages of Grounded Theory Analysis

Open coding: The researcher sliced the interview data and gave the same label to coded content with similar meanings. The coded data were continuously compared and grouped to extract subcodes (Chun Tie et al., 2019).

Axial coding: The researcher connects the subcodes established in the open coding at the level of meaning and generalises to a higher level of main categories (Chun Tie et al., 2019).

Selective coding: The researcher refines and integrates the main categories, clarifies their relationship with the core themes and finally builds a comprehensive conclusion framework. This process uses a continuous comparative analysis approach. The researchers move back and forth between stages to refine and integrate the categories and to clarify their relationships with the core themes, eventually forming the final conceptual framework. (Charmaz & Thornberg, 2021).

The data are considered saturated when the relationships between concepts are fully elucidated and no new information emerges, signifying that the analysis has reached its conclusion (Clarke et al., 2023).

4 Results

Using the 2I2A model as a guiding framework, we analysed the observed collaborative interactions to examine differences in the four collaborative Spaces and eight communication sub-dimensions across the LLM and non-LLMs groups.

4.1 Differences in collaboration spaces between two groups

Independent-sample t-tests were conducted for four collaboration spaces to examine differences between two groups with and without LLMs. Prior to the t-test, Levene's test for equality of variances was performed for all four collaboration spaces. For spaces where the assumption of equal variance was violated (Levene's test, $p < 0.05$), the Welch's t-test was employed, resulting in non-integer degrees of freedom (df). Conversely, for spaces where the equal variance assumption was met, the standard Student's t-test was used (df = 20).

Significant differences were found in the Identification Space and Interaction Space, while no significant differences were found in the Analysis and Accommodation Space (cf. **Table 4**). Given that four independent comparisons were performed, the Holm-Bonferroni correction (Holm, 1979) was applied to control the familywise error rate. After correction, the Identification Space (p = 0.0005) and Interaction Space (p = 0.0005) remained statistically significant. The Analysis Space (p = 0.467) and Accommodation Space (p = 0.937) remained non-significant. A detailed summary of the adjusted p-values for these four comparisons is provided in **Appendix 1**.

Table 4. Group Differences in Collaboration Spaces

Collaborative Space	Group	*Mean* (*SD*)	*T* (df)	*P*-value	Cohen's *d*
Identification Space	LLMs	18.64 (2.872)	-9.219	< 0.001	-3.930
	Non-LLMs	10.00 (1.183)	(13.297)		
Interaction Space	LLMs	28.18 (2.958)	-8.224	< 0.001	-3.510
	Non-LLMs	18.64 (2.462)	(20)		

Analysis Space	LLMs	23.55 (3.012)	-0.742	0.467	-0.316
	Non-LLMs	22.64 (2.730)	(20)		
Accommodation Space	LLMs	14.36 (2.838)	-0.080	0.937	-0.034
	Non-LLMs	14.27 (2.453)	(20)		

LLMs group (M = 18.64, SD = 2.872) collaborated significantly more frequently in Identification Space than the non-LLMs group (M = 10.00, SD = 1.183, $t(13.297) = -9.219$, $p < 0.001$, Cohen's d = -3.93). Similarly, in the Interaction Space, the LLMs group (M = 28.18, SD = 2.958) collaborated more frequently than the non-LLMs group (M = 18.64, SD = 2.462, $t\,(20) = -8.224$, $p < 0.001$, Cohen's d = -3.51). In contrast, no significant differences were found in the Analysis Space ($t(20) = -0.742$, $p = 0.467$, Cohen's d = -0.316). Similarly, in the Accommodation Space, the LLMs group (M = 14.36, SD = 2.838) and non-LLMs group (M = 14.27, SD = 2.453) also did not differ significantly ($t(20) = -0.080$, p = 0.937, Cohen's d = -0.034).

4.2 Differences in collaboration communications

Using MANOVA, Friedman's test, and the Wilcoxon Signed-Rank Test, significant differences were identified across the eight collaborative communication dimensions between the LLM and non-LLMs groups. Because eight follow-up comparisons were conducted across the communication dimensions, the Holm-Bonferroni correction (Holm, 1979) was applied to control the familywise error rate. All dimensions that were initially significant remained statistically significant after correction, while non-significant dimensions remained unchanged. A detailed summary of the adjusted p-values for all eight dimensions is provided in **Appendix 2**.

The LLMs group exhibited significantly higher frequencies in three dimensions related to idea and information generation. Specifically, for Information Perception, the LLMs group (M = 21.55, SD = 3.012) showed a substantially higher frequency compared to the non-LLMs group (M = 15.55, SD = 3.446), ($F(1, 20) = 18.906$, $p < 0.001$, partial $\eta^2 = 0.486$). The LLMs group also interacted more frequently for Information Selection and Accumulation (M = 21.55, SD = 4.741) than the non-LLMs group (M = 17.00, SD = 1.844), ($F(1, 20) = 8.784$, $p = 0.008$, partial $\eta^2 = 0.305$). Furthermore, expressions within the LLMs group were significantly more frequent for Structuring of Questions and Ideas (M = 49.27, SD = 4.941) compared to the non-LLMs group (M = 42.09, SD = 6.978), ($F(1, 20) = 7.761$, $p = 0.011$, partial $\eta^2 = 0.280$). Conversely, the non-LLMs group showed significantly higher frequencies in critical analysis and decision-making dimensions. For Evaluation, the non-LLMs group exhibited a significantly higher frequency (*M* = 35.36, *SD* = 5.500) compared to the LLMs group (*M* = 26.45, *SD* = 4.204), ($F(1, 20) = 18.217$, $p < 0.001$, partial $\eta^2 = 0.477$). Similarly, for Negotiation and Interpretation, expressions in the non-LLM condition occurred significantly more frequently (*M* = 44.36, *SD* = 8.488) than in the LLMs group (*M* = 34.73, *SD* = 4.819), ($F(1, 20) = 10.721$, $p = 0.004$, partial $\eta^2 = 0.349$). No significant differences were observed for Creative Output ($F(1, 20) = 1.879$, $p = 0.186$, partial $\eta^2 = 0.086$), with means of 11.18 (*SD* = 2.136) for the LLMs group and 12.73 (*SD* = 3.069) for the non-LLMs group.

The Wilcoxon Signed-Rank Test was used to analyse dimensions that did not meet normality assumptions. No significant differences were identified between the two groups for Cognitive Exchange ($Z = -1.512$, $p = 0.130$) or Solution Accommodation ($Z = -0.134$, $p = 0.893$). These results highlight the varying impact of LLMs across different collaborative dimensions, demonstrating strengths in perception, selection, and structuring but limitations in evaluation and negotiation.

4.3 Dynamic changes in collaboration at different stages

The study examined differences in the four collaboration spaces across early, middle, and later stages by groups with and without LLMs. **Figure 4** depicts the frequency change within each collaboration space. The figure reveals significant trends in collaboration dynamics across the stages.

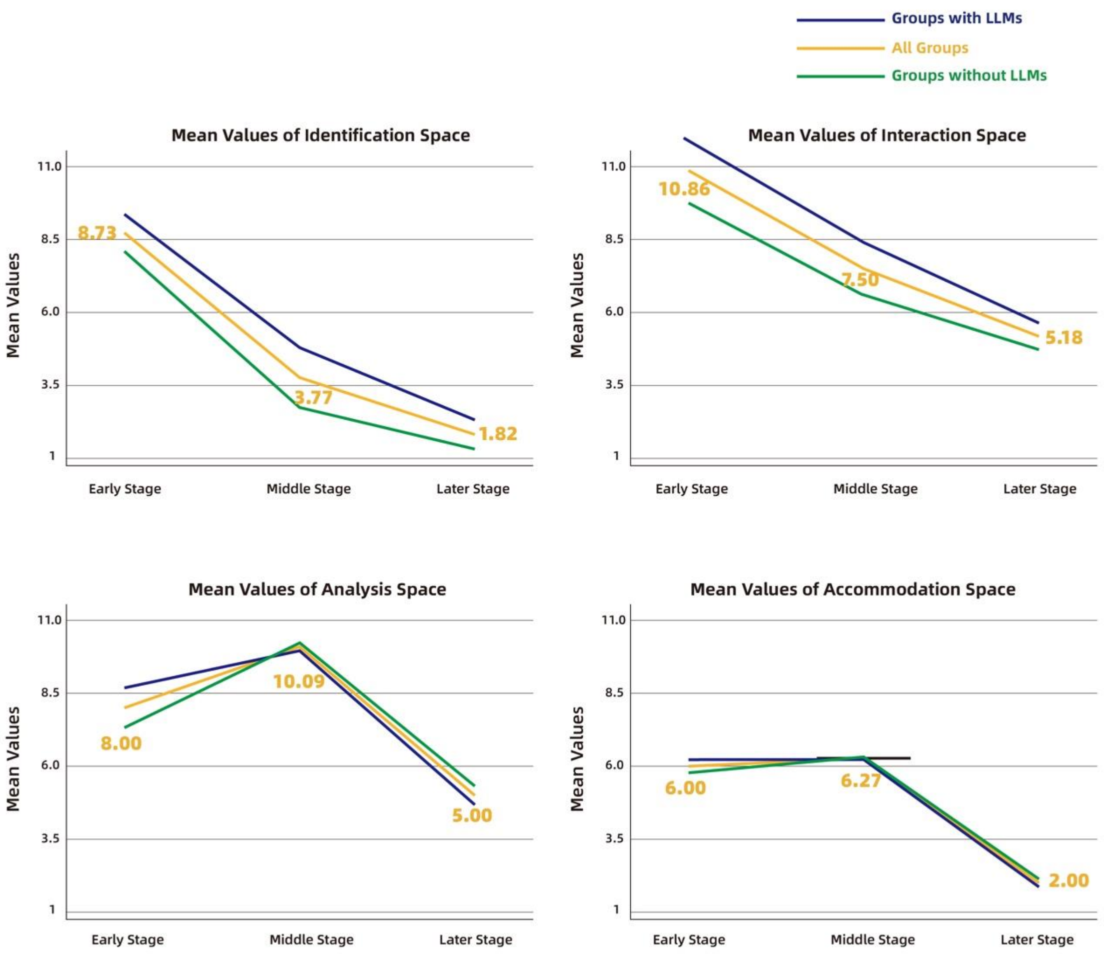

Figure 4. Changes in Mean Values of Collaboration Spaces Frequencies Across Stages

Differences in collaboration within Identification Space were significant across three stages ($\chi^2(2) = 33.86$, $p < 0.001$) based on the Friedman test. Post hoc Wilcoxon signed-rank tests with Bonferroni correction ($\alpha = 0.0167$) showed that there was significantly more collaboration in the early stages than in the middle and later stage ($p < 0.001$) and that collaboration in the middle stage was considerably more concentrated than in the later stage ($p = 0.002$). A detailed summary of the Bonferroni-adjusted post hoc comparisons for the Identification Space is provided in **Appendix 3.** This suggests a gradual decrease in the frequency of collaboration between collaborators in the Identification Space at all stages.

Collaboration within the Interaction Space showed a significant main effect of stage, ($F(1, 20) = 62.701$, $p < 0.001$, $\eta^2 = 0.758$). No significant quadratic trend was observed ($F(1, 20) = 0.590$, $p = 0.451$). The stage-group interaction effect was also non-significant ($F(1, 20) = 3.375$, $p = 0.081$), suggesting that the LLMs group did not differ significantly from the non-LLMs group in the pattern of collaboration trends over time. However, the LLMs group was consistently significantly higher than the non-LLMs group in the frequency of collaboration in the Interaction Space ($F(1, 20) = 68.290$, $p < 0.001$, $\eta^2 = 0.773$).

Collaboration within Analysis Space showed significant variation across stages, as indicated by a significant main effect ($F(1, 20) = 8.959$, $p = 0.007$, $\eta^2 = 0.309$). A significant quadratic trend was also observed ($F(1, 20) = 12.350$, $p = 0.002$, $\eta^2 = 0.382$), reflecting both linear and non-linear patterns of change. The interaction effect between stage and group was insignificant ($F(1.538, 30.767) = 1.894$, $p = 0.174$). Additionally, no significant group difference was observed in the overall frequency of collaboration in the Analysis Space ($F(1, 20) = 0.550$, $p = 0.467$, $\eta^2 = 0.027$).

Collaboration within Accommodation Space also differed significantly across the three stages ($\chi^2(2) = 16.949$, $p < 0.001$). Post hoc Wilcoxon signed-rank tests with Bonferroni correction ($\alpha = 0.0167$) showed that collaboration in the later stage was significantly fewer than in both the early stage and the middle stage ($p < 0.001$). At the same time, no significant difference was observed between the early and middle stages ($p = 0.793$). The corresponding Bonferroni-adjusted pairwise results for the Accommodation Space are summarised in **Appendix 4**. These findings suggest that collaboration of Accommodation Space remained stable between the early and middle stages but exhibited a substantial decline in the later stage.

4.4 Transition patterns under two groups

This section examines differences in collaborative spatial transitions between two groups to assess their transition patterns during Co-NPD (cf. **Table 5**). The results showed that LLM-involved groups exhibited higher mobility and complexity, reflected in higher overall chi-square values ($\chi^2(9) = 1435.584$, p < 0.001). In contrast, non-LLMs groups showed more fixed transition patterns ($\chi^2(9) = 799.073$, p < 0.001).

Table 5. Observed and Expected Frequencies of Collaborative Space Transitions under LLM and Non-LLM Conditions

Group	To Space				Total
	Accommodation	Analysis	Identification	Interaction	

Without LLMs	From Space	Accommodation	Count	0.0	133.0	23.0	0.0	156.0
			Expected	34.5	54.5	23.7	43.3	156.0
		Analysis	Count	124.0	0.0	0.0	121.0	245.0
			Expected	54.2	85.6	37.3	68.0	245.0
		Identification	Count	33.0	0.0	0.0	76.0	109.0
			Expected	24.1	38.1	16.6	30.2	109.0
		Interaction	Count	0.0	115.0	85.0	0.0	200.0
			Expected	44.2	69.9	34.0	55.5	200.0
	Total		Count	157.0	248.0	108.0	197.0	710.0
			Expected	157.0	248.0	108.0	197.0	710.0
With LLMs	From Space	Accommodation	Count	0.0	141.0	16.0	0.0	157.0
			Expected	26.9	43.9	34.4	51.8	157.0
		Analysis	Count	117.0	0.0	0.0	140.0	257.0
			Expected	44.0	71.9	56.3	84.7	257.0
		Identification	Count	41.0	0.0	0.0	164.0	205.0
			Expected	35.1	57.4	44.9	67.6	205.0
		Interaction	Count	0.0	117.0	186.0	0.0	303.0
			Expected	51.9	84.8	66.4	99.9	303.0
	Total		Count	158.0	258.0	202.0	304.0	922.0
			Expected	158.0	258.0	202.0	304.0	922.0

Under LLM conditions, the transition frequency between different collaborative spaces showed more substantial fluidity than expected. For example, the bidirectional transitions between the Identification Space and the Interaction Space were 164 and 186, markedly higher than their corresponding expected frequencies of 67.6 and 66.4, respectively. Similarly, the transitions between the Analysis Space and the Interaction Space were 140 and 117, also substantially elevated above their expected values.

In contrast, the overall transition frequencies of the non-LLMs group were significantly lower than those of the LLMs group. Specifically, the bidirectional transitions between the Identification Space and the Interaction Space were 76 and 85 in the non-LLMs group, markedly lower than the 164 and 186 observed under LLM conditions. Additionally, the transition from the Accommodation Space to the Analysis Space was 133 in the non-LLMs group, compared to 141 in the LLMs group.

On specific transitions, LLM participation did not significantly alter collaboration patterns. For example, transitions from the Interaction Space to the Analysis Space were 115 under non-LLMs conditions and 117 under LLM conditions. Transitions from the Analysis Space to the Accommodation Space were 124 under non-LLM conditions and 117 under LLM conditions. However, LLMs appeared to inhibit specific transitions; for example, transitions from the Accommodation Space to the Identification Space were only 16 under LLM conditions, lower than 23 under non-LLM conditions.

4.5 Grounded theory analysis: insights towards LLMs in collaborative design

The coding at open coding stage is explanatory, with researchers focusing on the relationships and variations between coding categories. **Figure 5** gives an initial encoding example.

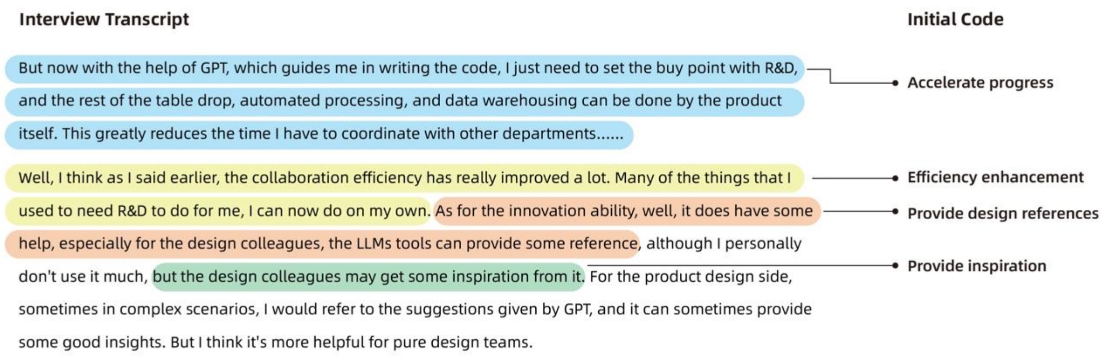

Figure 5. Example of Encoding in Open Coding Stage

After the 14th interview, the data were considered to be nearing saturation, as no new codes, categories, or conceptual insights emerged at this stage. To verify this judgement, the 15th and 16th interviews were analysed using the same open-coding procedure. Both transcripts failed to generate any additional codes or relationships, confirming that theoretical saturation had been reached according to established criteria that require consecutive interviews to yield no new conceptual contributions.

This stage ultimately obtained 125 preliminary codes, defined as 18 subcategories (cf. **Figure 6**). A detailed summary of the code count and category examples at each level is provided in **Table 6**.

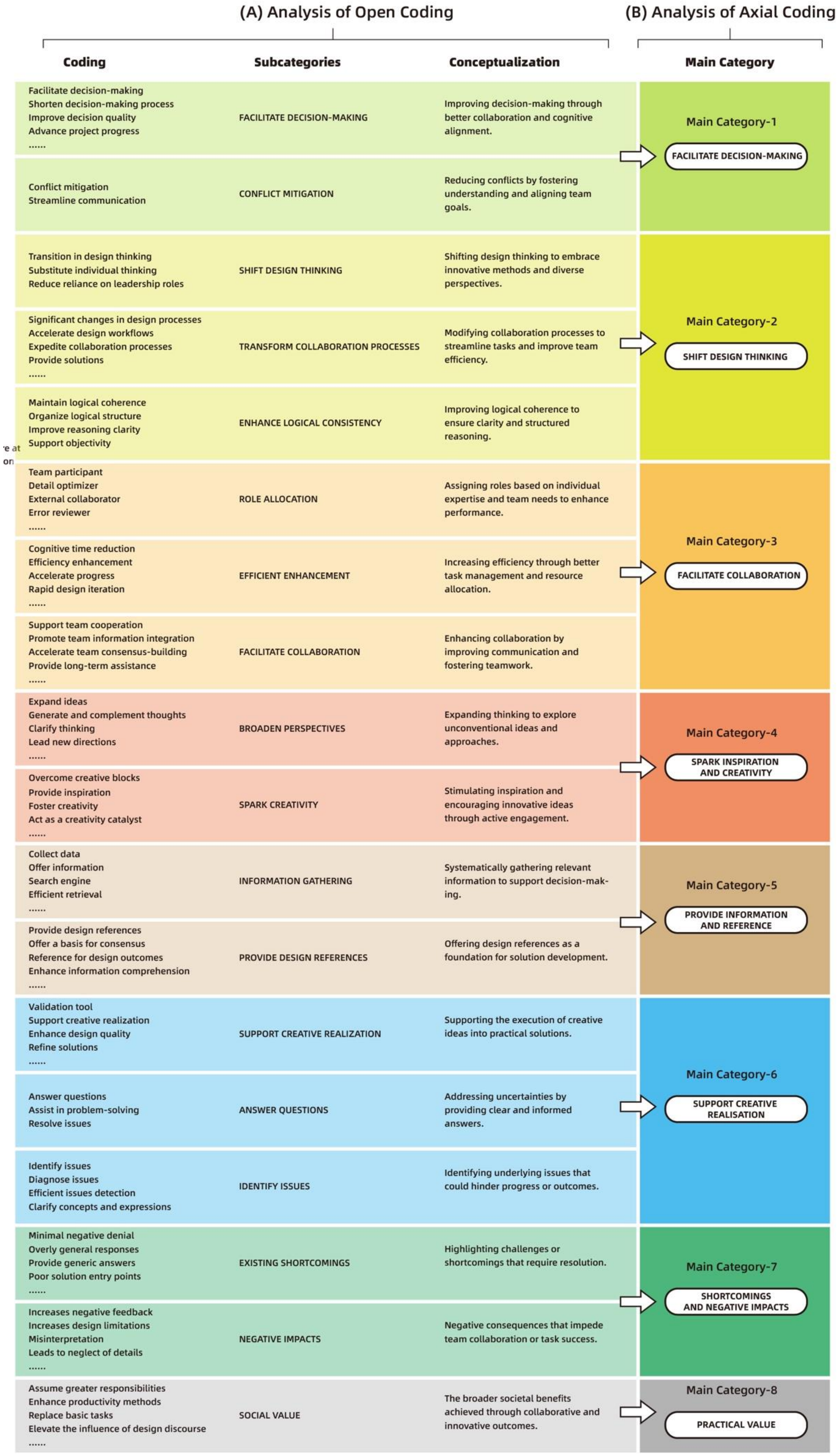
(A) Analysis of Open Coding
(B) Analysis of Axial Coding
Coding
Subcategories
Conceptualization
Main Category
Facilitate decision-making
Shorten decision-making process
Improve decision quality
Advance project progress
......
FACILITATE DECISION-MAKING
Improving decision-making through better collaboration and cognitive alignment.
Conflict mitigation
Streamline communication
CONFLICT MITIGATION
Reducing conflicts by fostering understanding and aligning team goals.
Main Category-1
FACILITATE DECISION-MAKING
Transition in design thinking
Substitute individual thinking
Reduce reliance on leadership roles
SHIFT DESIGN THINKING
Shifting design thinking to embrace innovative methods and diverse perspectives.
Significant changes in design processes
Accelerate design workflows
Expedite collaboration processes
Provide solutions
......
TRANSFORM COLLABORATION PROCESSES
Modifying collaboration processes to streamline tasks and improve team efficiency.
Maintain logical coherence
Organize logical structure
Improve reasoning clarity
Support objectivity
ENHANCE LOGICAL CONSISTENCY
Improving logical coherence to ensure clarity and structured reasoning.
Main Category-2
SHIFT DESIGN THINKING
Team participant
Detail optimizer
External collaborator
Error reviewer
......
ROLE ALLOCATION
Assigning roles based on individual expertise and team needs to enhance performance.
Cognitive time reduction
Efficiency enhancement
Accelerate progress
Rapid design iteration
......
EFFICIENT ENHANCEMENT
Increasing efficiency through better task management and resource allocation.
Support team cooperation
Promote team information integration
Accelerate team consensus-building
Provide long-term assistance
......
FACILITATE COLLABORATION
Enhancing collaboration by improving communication and fostering teamwork.
Main Category-3
FACILITATE COLLABORATION
Expand ideas
Generate and complement thoughts
Clarify thinking
Lead new directions
......
BROADEN PERSPECTIVES
Expanding thinking to explore unconventional ideas and approaches.
Overcome creative blocks
Provide inspiration
Foster creativity
Act as a creativity catalyst
......
SPARK CREATIVITY
Stimulating inspiration and encouraging innovative ideas through active engagement.
Main Category-4
SPARK INSPIRATION AND CREATIVITY
Collect data
Offer information
Search engine
Efficient retrieval
......
INFORMATION GATHERING
Systematically gathering relevant information to support decision-making.
Provide design references
Offer a basis for consensus
Reference for design outcomes
Enhance information comprehension
......
PROVIDE DESIGN REFERENCES
Offering design references as a foundation for solution development.
Main Category-5
PROVIDE INFORMATION AND REFERENCE
Validation tool
Support creative realization
Enhance design quality
Refine solutions
......
SUPPORT CREATIVE REALIZATION
Supporting the execution of creative ideas into practical solutions.
Answer questions
Assist in problem-solving
Resolve issues
ANSWER QUESTIONS
Addressing uncertainties by providing clear and informed answers.
Identify issues
Diagnose issues
Efficient issues detection
Clarify concepts and expressions
IDENTIFY ISSUES
Identifying underlying issues that could hinder progress or outcomes.
Main Category-6
SUPPORT CREATIVE REALISATION
Minimal negative denial
Overly general responses
Provide generic answers
Poor solution entry points
......
EXISTING SHORTCOMINGS
Highlighting challenges or shortcomings that require resolution.
Increases negative feedback
Increases design limitations
Misinterpretation
Leads to neglect of details
......
NEGATIVE IMPACTS
Negative consequences that impede team collaboration or task success.
Main Category-7
SHORTCOMINGS AND NEGATIVE IMPACTS
Assume greater responsibilities
Enhance productivity methods
Replace basic tasks
Elevate the influence of design discourse
......
SOCIAL VALUE
The broader societal benefits achieved through collaborative and innovative outcomes.
Main Category-8
PRACTICAL VALUE

Figure 6. Grounded Theory Coding Process Flow

Table 6. Summary of Grounded Theory Core Categories

Coding Level	Item Count	Representative Subcategory	Supporting Quotes
Level 1: Open Coding	125 initial codes	Improve decision quality; Conflict mitigation; Error reviewer; Lead new directions	Conflict mitigation: "Right, the conflicts went down. When people have GPT acting like an 'expert' giving suggestions, they feel more confident and have a clearer sense of which option is more workable."
Level 2: Axial Coding	18 subcategories	Facilitate decision-making; Role allocation; Shift Design thinking; Transform collaboration processes	Transform collaboration processes: "The design process did improve. For example, when generating a 3D model, I would first use images to generate prompts, and then..."
Level 3: Core Category	8 core themes	Facilitate decision-making	"So in a way, it (GPT) basically ends up playing a sort of decision-maker role."
		Shift design thinking	"With GPT in the mix, designers' responsibilities have expanded, and their way of thinking has shifted as well."
		Facilitate collaboration	"The collaboration went more smoothly because communication became more fluid, and team members could understand each other's ideas more easily."
		Spark inspiration and creativity	"For individuals, AI support can boost their creativity."
		Provide information and reference	"In the early stages, these tools mainly helped me with searching and research, giving me access to the relevant knowledge and information I needed."
		Support creative realisation	"It (GPT) really did help us bring the design to life and carry out the solution more effectively."
		Shortcomings and negative impacts	"The downside is that for some of the newest knowledge, GPT might not have the latest information, so I still need to use traditional search tools in those cases."
		Practical value	"But it may create more AI-related career

opportunities and broaden designers' career paths."

During the Axial Coding stage, the subcategories identified during open coding are further integrated into a coherent structure and constructed into abstract conceptual classifications based on descriptive details to explore the interrelationships and differences among these categories (Clarke et al., 2023). Finally, 18 subcategories were refined into eight categories, including Facilitate Decision-making, Shift Design Thinking, Facilitate Collaboration, Spark Inspiration and Creativity, Provide Information and Reference, Support Creativity Realisation, Shortcomings and Negative Impacts, and Social Value (**cf. Figure 6**).

During the Selective Coding stage, researchers refine and integrate the main categories extracted based on axial coding, identifying the core categories and their relationships with other categories. During this process, researchers connect all categories related to the core category to construct a comprehensive theoretical framework. "Practical Value" was identified as the core category connecting all other themes. Based on this, we developed a conceptual model that illustrates the perceived roles and practical contributions of LLMs in supporting collaborative new product development (cf. **Figure 7**).

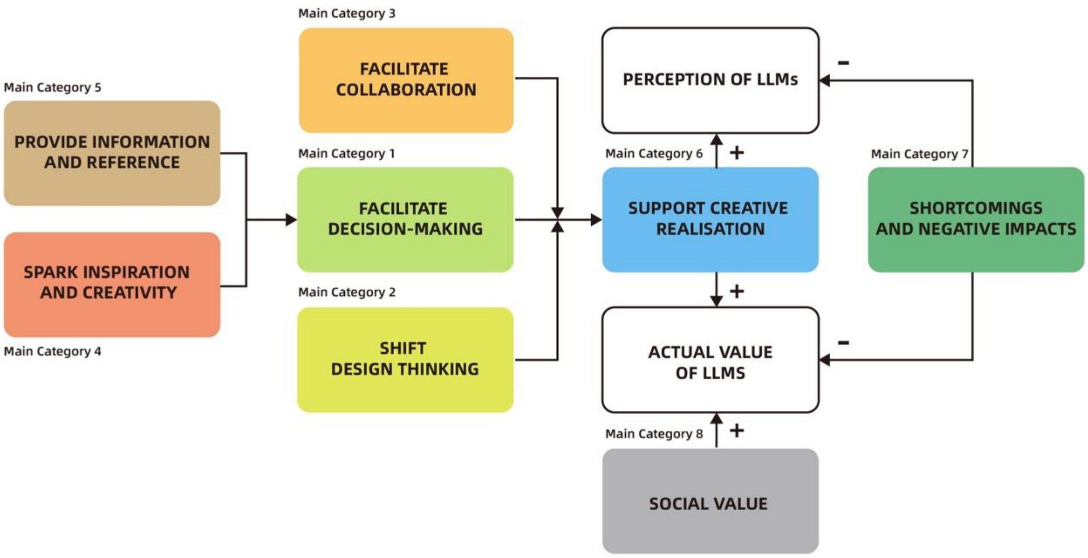

Figure 7. The Relationship Model Diagram between the perceptions and Actual Value of LLMs during Co-NPD of Collaborators

5 Discussion

5.1 Promotion and limitations of LLMs on collective cognition

This paper provides insights into the role of LLMs during Co-NPD on collective cognition, aiming to explicitly address RQ1. The experimental results indicate that teams supported by LLMs collaborated significantly more frequently than non-LLM teams in both the Identification and Interaction Spaces. This effect was particularly pronounced in the sub-dimensions of Information Perception, Information Selection and Accumulation, and the Structuring of Questions and Ideas. According to the 2I2A model, the Identification Space represents the initial stage of shared perception and knowledge input for collective cognition, and the Interaction Space represents the core process of knowledge coordination, structuring, and preliminary integration.

Grounded theory analysis further indicates that LLMs exhibit a distinctive capability in group information processing. Themes such as "Facilitate collaboration" and "Provide information and reference" (see Section 4.5) help explain the quantitative pattern. Several interviewees noted that LLMs were able to organise and present large volumes of complex information directly within the communication flow, thereby accelerating the team's shared understanding of the design problem. For instance, interviewee 16 emphasised this inspirational function of LLMs, stating,

> "*In the early design stage, our team relied heavily on large language models to help us gather information.*"

Similarly, interviewee 1 noted the efficiency benefits, explaining,

> "*At the beginning, we needed to conduct a lot of background research, and using GPT saved us some of that time and accelerated the design process.*"

Taken together, these accounts indicate that LLMs reduce early-stage search and filtering effort by providing structured, task-relevant information in a timely manner. This support enables collaborators to identify relevant content more frequently and to organise it more effectively into problem statements and sub-tasks, resulting in increased activity in the Identification and Interaction Spaces. This mechanism aligns with established literature on collective cognition and knowledge management, which emphasises that structured and well-timed information flows are essential for forming shared mental models (Mathieu et al., 2000), Accordingly, LLMs' ability to streamline the creation and circulation of structured inputs plays a meaningful role in accelerating the initial formation of collective cognition.

While LLMs demonstrate clear advantages in these early cognitive stages, their support does not extend equally to the more complex analytical processes required in the later stages of collaboration. Collaboration in the Analysis and Accommodation Spaces typically requires logical deduction, complex decision-making, and the integration of diverse perspectives. The quantitative results show no significant differences between LLM and non-LLMs groups in these Spaces, and the qualitative findings echo this pattern. Under the theme "Shortcomings and Negative Impacts", interviewees frequently noted accuracy issues, limited contextual reasoning, and a tendency to offer surface-level suggestions. These limitations collectively constrain the utility of LLMs during deeper analytical reasoning. For instance, interviewee 1 highlighted this lack of critical challenge in LLM interactions, stating,

> "*Because it (GPT) always follows along with my line of thinking and lacks the kind of questioning that a human teammate would offer, I can't help worrying that the design might become limited.*"

Similarly, interviewee 9 commented on the inefficiency that may arise during use, explaining,

> "*GPT can't always fully understand what I input, so I have to keep adjusting and re-entering information, which takes time.*"

Furthermore, interview analyses indicated that collaborators often preferred to rely on their professional judgement and direct discussion rather than on AI-generated suggestions when engaging in evaluative or interpretive tasks. Non-LLM teams performed better in sub-dimensions such as Evaluation and Negotiation and Interpretation. Interviewee 3 remarked that AI-generated ideas sometimes narrowed their thinking, noting that

> "*In some steps, the LLMs replaced my own thinking, and it really somehow interfered with my design process.*"

In summary, LLMs offer clear benefits in facilitating early cognitive alignment and structuring collaborative problem spaces. However, their contribution becomes more limited in collaborative contexts that require interpretive depth, nuanced negotiation, and complex integrative reasoning (Cuskley et al., 2024).

5.2 The role of LLMs in collaborative cognitive mobility

The results reveal the profound effect of LLMs on the dynamic characteristics of group collaboration during Co-NPD. Across the three stages of the experiment, the LLMs group consistently demonstrated a higher frequency of cooperation in the Identification and Interaction Spaces compared to the non-LLMs group. However, both groups exhibited a decreasing trend in collaboration activities over time. This pattern suggests that as collaboration progressed, the team's focus gradually shifted from initial problem identification and information exchange toward more complex tasks, such as analysis and adaptation. The consistently higher frequencies in the Identification and Interaction Spaces for the LLMs group indicate that LLMs help teams work more efficiently in early problem identification and information sharing. These advantages accelerate the development of shared cognition. As a result, teams demonstrate greater cognitive flexibility in the early stages of collaboration.

Collaboration patterns in the Analysis and Accommodation Spaces were more complex. Collaboration in the Analysis Space did not show significant differences between the two groups, with activity peaking in the middle stage before declining in the later stage. This suggests that activities in the Analysis Space are mainly concentrated in the middle stage when the team needs to conduct an in-depth assessment and trade-offs of the collected information and preliminary options. The absence of group differences indicates that teams relied primarily on members' expertise and deductive reasoning rather than LLM assistance in this Space. In the Accommodation Space, collaboration remained stable during the early and middle stages but dropped sharply in the later stage. This pattern reflects that fewer integrative actions occur once teams converge toward a tentative solution.

The result of transitions in the collaboration space showed that the significantly higher fluidity and complexity brought by LLMs is associated with a shift in the team's collaboration pattern (as detailed in **Table 5**). This change indicates a transition towards a more iterative and agile approach, indicating that LLMs actively reform the team's collaborative dynamics. LLMs significantly enhanced collaboration dynamics, especially in the bidirectional transitions between the Identification Space and the Interaction Space, as well as the switch from the Analysis Space to the Interaction Space, facilitating the team's information exchange and cognitive iteration. This increase in fluidity is closely linked to the qualitative themes "Support Creative Realisation" and "Facilitate Collaboration," where LLMs provide instant feedback, generate rapid prototypes, and offer structured responses. This efficiency was frequently noted by participants, for example:

Interviewee 7 described this efficiency vividly, noting that

> *"GPT acted like a kind of conceptual design partner. It mainly helped propose ideas and concepts, giving us initial directions and references for the design."*

Similarly, interviewee 2 highlighted its role in expanding the team's thinking:

> *"LLMs can also broaden my imagination by presenting knowledge that I might not be familiar with."*

This tool-supported agility allows teams to cycle faster between identifying needs and structuring options, thereby enhancing overall cognitive mobility. However, the involvement of LLMs did not significantly alter the frequency of transitions in specific pathways (such as Interaction Space to Analysis Space, Analysis Space to Accommodation Space). A particularly notable feature is the low frequency of transitions from the Accommodation Space back to the Identification Space. This pattern suggests a collaboration flow that relies less on cyclical feedback loops and problem re-definition once a solution has begun to take shape. This structural characteristic has a meaningful influence on the overall collaboration pattern.

Despite the clear improvements in collaboration fluidity, LLM use also introduces potential drawbacks. For example, frequently switching collaboration spaces may interfere with the team's deeper thinking about critical tasks (interviewee 7). At the same time, over-reliance on technology may weaken the team's long-term programme development continuity (Interviewee 8). Therefore, the practical application of LLMs must be tailored to different task types. The team should design an optimised strategy that promotes collaborative flexibility while retaining the ability to analyse in depth and iterate on innovation.

5.3 Practical application value of LLMs during Co-NPD

The Relationship Model Diagram in section 4.5 (cf. **Figure 6**) reveals the practical application value of LLMs in Co-NPD, thereby addressing RQ3. Research has found that actual value, as the core category, is closely related to all other main categories, especially those related to "support creative realisation", "shortcomings and negative impacts", and "social value". Among them, "support creative realisation" has a positive driving effect on actual value, while "shortcomings and negative impacts" have a negative effect.

"Social value" reflects the extensive influence of LLMs at the team and social levels. These qualitative findings not only highlight the practical impact (RQ3) but also provide the underlying mechanistic framework used to interpret the quantitative differences observed across the collaborative spaces (as discussed in 5.1 and 5.2).

Support creative realisation

Multiple interviewees (Interviewees 1, 3, 5, 7, 8, 9, 10, 11, 12, 13, 14) mentioned that LLMs play a significant role in creative generation and concept optimisation. Interviewee 13 pointed out, "GPT sometimes provides very inspiring ideas, and I tend to see it as a motivator." Interviewee 14 pointed out: "In the middle stage of design, it can solve operational problems, which is more effective than search engines. In the later stage, it is mainly used for debugging and optimisation. These tools allow designers to take on more responsibilities, such as modelling to final rendering and optimisation work, reducing dependence on programmers. Therefore, the role of designers becomes more diverse, and the workflow becomes more efficient". Interviewee 11 mentioned, "For innovation capabilities, AI provides a wide range of inspiration sources not limited to information cocoons. Through LLMs, we can gain more diverse perspectives and creativity."

Based on the interviews and protocol analysis results, LLMs have significantly improved the team's innovation ability and task efficiency, especially in the early stages of collaboration, thanks to their inspiring and diverse perspectives. For example, in Identification and Interaction spaces, LLMs can accelerate the establishment of collective cognition by providing inspiration and optimizing information flow, promoting team formation of shared perception and intention. Interviewees consistently described that LLMs helped them "see more possibilities" (interviewee 15), "organise messy information" (interviewee 6), and "quickly form initial ideas" (interviewee 13). These patterns help explain the mechanism underlying the statistical differences. LLMs broaden the accessible information base and enable faster organisation of that information. This process leads to the higher frequencies of Identification and Interaction activities observed in the experiment.

Disadvantages and negative impacts

Some interviewees mentioned that LLMs can lead to collaborators' overreliance on technology, thus limiting their creativity. Interviewee 8 mentioned, "*LLM tool can be very helpful, but it may not be possible to over-rely on it, (as it) may limit the output of some of our ideas.*" Interviewee 13 notes that "*It may limit the viewpoint to the GPT idea.*" This suggests that such interventions not only diminish the designer's initiative but also hurt the naturalness of the collaborative development process.

Furthermore, there are limitations in the technical capabilities of LLMs. Interviewee 3 mentioned, "*I found that the LLMs replaced my thinking in some steps, which affected my design process.*" Interviewee 16 mentions that "*There is still a lot of room for improvement in the accuracy of information gathering in LLMs*." This suggests that although the model provides information efficiently, its reliability is inconsistent. As a result, designers must spend additional effort on verification, which may affect overall workflow efficiency.

Collectively, LLMs' negative impacts are mainly in technology dependency, lack of information accuracy, and weakened collaborative naturalness. These issues remind us that when LLMs are introduced to participate in Co-NPD, their application scenarios and usage should be reasonably planned to ensure that the technology's advantages are not exploited at the cost of team creativity and collaboration quality.

These qualitative limitations also help interpret the quantitative findings. In later-stage collaborative spaces that require negotiation, integration and evaluation, interviewees reported hesitation or distrust toward LLM-generated content. This aligns with the lower frequencies observed in the Analysis and Accommodation Spaces. Thus, the qualitative insights clarify why LLMs enhanced early-stage cognition but did not extend their benefits to deeper collaborative reasoning.

Finally, beyond these technical and cognitive limitations, LLM deployment in Co-NPD also raises ethical considerations. Issues such as data privacy, confidentiality of proprietary design information, and potential algorithmic biases in AI-generated suggestions may affect team trust and collaborative quality. These findings highlight the importance of carefully planning how LLMs are used. It is also essential to implement safeguards such as secure data handling and critical evaluation of AI outputs to mitigate risks while supporting team creativity.

Social value

The feedback from interviewees indicates that the social value of LLMs in Co-NPD is mainly reflected in optimising team power structure, transforming foundational work, and improving individual abilities. Interviewee 16 pointed out, “This (LLM tool) also gives product managers greater say in the team, which may be the most crucial point: the maximum value brought by LLMs.” This indicates that LLMs optimise the organisation’s power structure and empower specific core roles with more significant influence in the collaborative process.

The social value of LLMs is also reflected in promoting the automation and digital transformation of specific foundational work. Interviewee 2 pointed out, “It is very likely to replace some basic work.” In contrast, Interviewee 9 pointed out that “It reduces the demand for basic work roles, while the demand for roles in the creative and decision-making stages will increase.” This indicates that the involvement of LLMs enables the automation of some simple and repetitive work through Co-NPD, freeing up more resources and energy for high-value creative and decision-making work, and can promote the trend of developers transitioning from fundamental work to higher-level skills. In addition, LLMs significantly affect the allocation of responsibilities within the team, as mentioned in Interviewee 9: “These tools allow designers to take on more responsibilities, such as modelling to final rendering and optimisation work, reducing reliance on programmers.” Interviewee 7 added, “LLMs reduce the workload of planning and leadership roles.”

In summary, the practical application value of LLMs during Co-NPD mainly lies in creative support, process optimisation, and social impact. Although LLMs still face technological dependence and information accuracy challenges, they demonstrate enormous potential in empowering team innovation, improving task efficiency, and driving organisational role transformation.

6 Conclusion

6.1 Summary of contributions

This study provides an in-depth examination of the influence of LLMs during Co-NPD and makes three primary contributions.

First, it develops and validates the 2I2A model from a collective cognition perspective.

The model systematically captures cognitive flow and spatial transitions by organising team interactions into four collaboration spaces and eight sub-dimensions. This framework offers a novel analytical lens for understanding the dynamic cognitive processes involved in Co-NPD. It also supports more precise evaluation and optimisation of complex collaborative activities.

Second, through a quasi-experimental design, this study systematically evaluates the role of LLMs across different stages of Co-NPD. LLMs significantly enhanced collaboration efficiency in the Identification and Interaction Spaces, particularly in Information Perception and Structuring of Questions and Ideas during the early phase. These findings demonstrate that LLMs substantially facilitate the early construction and integration of collective cognition.

However, in the Analysis and Accommodation Spaces, collaborators relied more heavily on their own experience and knowledge. In these contexts, LLMs did not provide significant facilitation. In addition, LLMs increased the overall fluidity of transitions between most collaborative spaces, indicating a shift toward a more iterative and agile collaboration pattern (addressing RQ2). Yet, the transition from the Accommodation Space back to the Identification Space showed limited or even inhibitory effects. This pattern suggests boundaries in LLMs' ability to support cyclical reinterpretation and problem reframing. Overall, while LLMs optimise collaboration dynamics and enhance cognitive mobility, their value is more restricted in tasks requiring deep analysis or complex decision-making.

Finally, the paper supplements the analysis with qualitative Grounded Theory, revealing the practical value and limitations of LLMs in Co-NPD. LLMs offer strong support for idea generation**,** process optimisation**,** and expansion of cognitive breadth, thereby improving innovation efficiency within teams. Moreover, they contribute to the automation of repetitive tasks within the collaborative process. This shift also fosters role diversification, enabling team members to assume broader responsibilities. However, LLMs may also introduce drawbacks, including overreliance on AI-generated suggestions, reduced creative exploration, and a diminished depth of human–human collaboration. Their inconsistent accuracy and limited contextual reasoning often require teams to manually verify AI-generated content. This extra verification reduces efficiency and weakens trust, particularly in later-stage analytical and integrative tasks.

6.2 Limitations

This study has several limitations. First, the experiment was conducted in a controlled environment with participants drawn from academic backgrounds. The relatively small sample size of 44 participants (22 collaborative teams) may restrict the statistical generalisability of the findings and may not fully capture the complexity of Co-NPD practices in real organisational settings. This limits the broader applicability of the conclusions.

Second, as LLM technologies continue to evolve rapidly, their capabilities are likely to expand. Future, more advanced models may therefore exhibit behaviours or affordances that fall beyond the scope of those examined in this study. In addition, participants in the LLM-assisted condition were allowed to choose between two commonly used models (ChatGPT-4.0 and Kimi) during the Co-NPD tasks, which may introduce minor variability across teams. While this reflects real-world Co-NPD settings where different teams may adopt different LLM tools, we acknowledge that this flexibility may introduce variability. Future studies may therefore consider standardising the LLM model to further reduce potential confounds.

Third, the study focuses on immediate, short-term collaboration dynamics and does not investigate the effects of LLMs in long-term collaboration or in extended iterative design cycles. The absence of longitudinal evidence may limit understanding of cumulative or delayed impacts.

Crucially, while our qualitative data touched on negative impacts, the study does not provide a systematic quantitative or qualitative examination of these issues. The analysis does not cover key ethical and governance challenges associated with LLM deployment. These challenges include issues such as data privacy, the confidentiality of proprietary design information, and the influence of algorithmic biases on design outcomes. Addressing these issues will be essential for developing comprehensive and responsible frameworks for AI-supported Co-NPD.

6.3 Future research directions

Building upon the limitations identified in this study and the current findings, we propose several concrete directions for future research to deepen the understanding of LLM impacts on collaborative cognition.

First, exploring diverse contexts and professional settings. Future work should involve larger sample sizes and include professional designers or engineers in real-world industry environments. Additionally, research should investigate LLM use in multicultural and multidisciplinary teams. This will help validate the applicability of the 2I2A model and examine LLM effects under practical constraints such as time pressure and budget limitations.

Second, conducting longitudinal studies. Future research should adopt longitudinal designs to investigate the long-term impact of continuous LLM use on team creativity, skill development, and the potential erosion of independent critical thinking and design expertise.

Third, quantifying negative impacts and establishing responsible guidelines. Future studies should develop specific metrics and protocols to assess potential negative socio-cognitive effects of LLMs. This assessment must cover critical areas such as technological dependency, shifts in communication equity, algorithmic bias, and risks to data confidentiality and intellectual property in Co-NPD. This comprehensive approach will ensure a more balanced and rigorous assessment of both the benefits and drawbacks of LLM integration. Building upon these findings, future work should propose and validate concrete guidelines and governance frameworks to ensure the responsible and ethical application of LLMs in collaborative design processes.

Finally, conducting industry-specific case studies and external validity tests. Future work should incorporate industry-specific case studies to examine the 2I2A model and LLM effects in concrete industrial settings, such as design studios or technology firms. This will allow rigorous testing of applicability under real-world constraints, including varying timelines, resource limitations, and high project stakes.

Overall, the 2I2A model presented in this study is an initial step. Future research should continue to validate and refine the model across diverse teams, industrial contexts, and long-term collaborative processes.

Competing interests

The author(s) declare no competing interests.

Data availability

The datasets generated and analysed during this study contain human interaction data and are subject to ethical approval conditions, participant consent, and applicable privacy and data-protection requirements.

An anonymised subset of the data, including segment-level coded datasets, coding schemes, grounded theory analysis outputs, and quantitative analysis files (SPSS), has been made available to the editors and reviewers during peer review via the journal's submission system.

Original audio recordings and full verbatim conversational transcripts cannot be publicly shared due to ethical and confidentiality constraints, as they contain potentially identifiable interpersonal communication data and were not consented for unrestricted disclosure.

Further anonymised data supporting the findings of this study may be made available from the corresponding author upon reasonable request, subject to ethical approval and participant consent.

Funding declaration

The authors declare that no funding was received for this study.

Ethical approval statement

This study involved human participants through interviews and quasi-experimental collaborative design tasks. No biomedical, clinical, or invasive procedures were conducted. All data were collected anonymously and handled confidentially for research purposes only. The study was conducted in accordance with the Declaration of Helsinki and institutional guidelines for research involving human participants. Ethical approval was granted by the Zhejiang University Ethics Committee (Approval Date: 30 April 2024; the committee does not issue approval numbers).

Informed consent

All participants provided written informed consent prior to taking part in the study, between 14 May and 30 July 2024. They were informed of the study purpose, procedures, potential risks, and their right to withdraw at any time without penalty. No personally identifiable or sensitive information was collected, and all audio recordings and transcripts were anonymised during analysis. All procedures adhered to institutional ethical standards for research involving human participants.

Reference

Alaskar, T. H., Alsadi, A. K., Aloulou, W. J., & Ayadi, F. M. (2024). Big Data Analytics, Strategic Capabilities, and Innovation Performance: Mediation Approach of Organizational Ambidexterity. *Sustainability*, *16*(12), 5111. https://doi.org/10.3390/su16125111

Alavi, M., & Leidner, D. E. (2001). Review: Knowledge Management and Knowledge Management

Systems: Conceptual Foundations and Research Issues. *MIS Quarterly*, *25*(1), 107–136. https://doi.org/10.2307/3250961

Alkhwaldi, A. F. (2024). Understanding the acceptance of business intelligence from healthcare professionals' perspective: An empirical study of healthcare organizations. *International Journal of Organizational Analysis*, *32*(9), 2135–2163. https://doi.org/10.1108/IJOA-10-2023-4063

Alkhwaldi, A. F., Abdulmuhsin, A. A., Masa'deh, R., & Abu-AlSondos, I. A. (2025). Business intelligence adoption in higher education: The role of data-driven decision-making culture and UTAUT. *Journal of International Education in Business*, *18*(4), 484–504. https://doi.org/10.1108/JIEB-08-2024-0110

Andrews, R. W., Lilly, J. M., Srivastava, D., & Feigh, K. M. (2023). The role of shared mental models in human-AI teams: A theoretical review. *Theoretical Issues in Ergonomics Science*, *24*(2), 129–175. https://doi.org/10.1080/1463922X.2022.2061080

Asarnow, S. (2020). Shared Agency Without Shared Intention. *The Philosophical Quarterly*, *70*(281), 665–688. https://doi.org/10.1093/pq/pqaa012

Badke-Schaub, P., Neumann, A., Lauche, K., & Mohammed, S. (2007). Mental models in design teams: A valid approach to performance in design collaboration? *CoDesign*, *3*(1), 5–20. https://doi.org/10.1080/15710880601170768

Ball, L. J., Ormerod, T. C., & Morley, N. J. (2004). Spontaneous analogising in engineering design: A comparative analysis of experts and novices. *Design Studies*, *25*(5), 495–508. https://doi.org/10.1016/j.destud.2004.05.004

Beverland, M. B., Micheli, P., & Farrelly, F. J. (2016). Resourceful Sensemaking: Overcoming Barriers between Marketing and Design in NPD. *Journal of Product Innovation Management*, *33*(5), 628–648. https://doi.org/10.1111/jpim.12313

Bowen, G. A. (2008). Naturalistic inquiry and the saturation concept: A research note. *Qualitative Research*, *8*(1), 137–152. https://doi.org/10.1177/1468794107085301

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*, *8*(9), 1643–1655. https://doi.org/10.1038/s41562-024-01959-9

Carragher, D. J., Sturman, D., & Hancock, P. J. B. (2024). Trust in automation and the accuracy of human–algorithm teams performing one-to-one face matching tasks. *Cognitive Research: Principles and Implications*, *9*(1), 41. https://doi.org/10.1186/s41235-024-00564-8

Carraro, M., Furlan, A., & Netland, T. (2025). Unlocking team performance: How shared mental models drive proactive problem-solving. *Human Relations*, *78*(4), 407–437. https://doi.org/10.1177/00187267241247962

Carroll, J. M. (1997). Human–computer interaction: Psychology as a science of design. *International Journal of Human-Computer Studies*, *46*(4), 501–522. https://doi.org/10.1006/ijhc.1996.0101

Cha, K. J., Kim, Y. S., Park, B., & Lee, C. K. (2015). Knowledge Management Technologies for Collaborative Intelligence: A Study of Case Company in Korea. *International Journal of Distributed Sensor Networks*, *11*(9), 368273. https://doi.org/10.1155/2015/368273

Chalmers, P. A. (2003). The role of cognitive theory in human–computer interface. *Computers in Human Behavior*, *19*(5), 593–607. https://doi.org/10.1016/S0747-5632(02)00086-9

Chan, C.-S. (1990). Cognitive processes in architectural design problem solving. *Design Studies*, *11*(2), 60–80. https://doi.org/10.1016/0142-694X(90)90021-4

Charmaz, K., & Thornberg, R. (2021). The pursuit of quality in grounded theory. *Qualitative Research in Psychology*, *18*(3), 305–327. https://doi.org/10.1080/14780887.2020.1780357

Charters, E. (2003). The Use of Think-aloud Methods in Qualitative Research An Introduction to Think-aloud Methods. *Brock Education Journal*, *12*(2), Article 2. https://doi.org/10.26522/brocked.v12i2.38

Chen, O., Paas, F., & Sweller, J. (2023). A Cognitive Load Theory Approach to Defining and Measuring Task Complexity Through Element Interactivity. *Educational Psychology Review*, *35*(2), 63. https://doi.org/10.1007/s10648-023-09782-w

Chen, Y., Wang, Y., Nevo, S., Benitez-Amado, J., & Kou, G. (2015). IT capabilities and product innovation performance: The roles of corporate entrepreneurship and competitive intensity. *Information & Management*, *52*(6), 643–657. https://doi.org/10.1016/j.im.2015.05.003

Christiansen, J. K., & Varnes, C. J. (2009). Formal Rules in Product Development: Sensemaking of Structured Approaches. *Journal of Product Innovation Management*, *26*(5), 502–519. https://doi.org/10.1111/j.1540-5885.2009.00677.x

Chun Tie, Y., Birks, M., & Francis, K. (2019). Grounded theory research: A design framework for novice researchers. *SAGE Open Medicine*, *7*, 2050312118822927. https://doi.org/10.1177/2050312118822927

Clarke, A., Healy, K., Lynch, D., & Featherstone, G. (2023). The Use of a Constructivist Grounded Theory Method—A Good Fit for Social Work Research. *International Journal of Qualitative Methods*, *22*, 16094069231186257. https://doi.org/10.1177/16094069231186257

Cody, W. F., Kreulen, J. T., Krishna, V., & Spangler, W. S. (2002). The integration of business intelligence and knowledge management. *IBM Systems Journal*, *41*(4), 697–713. https://doi.org/10.1147/sj.414.0697

Cohn, C., Snyder, C., Montenegro, J., & Biswas, G. (2024). Towards a Human-in-the-Loop LLM Approach to Collaborative Discourse Analysis. In A. M. Olney, I.-A. Chounta, Z. Liu, O. C. Santos, & I. I. Bittencourt (Eds.), *Artificial Intelligence in Education. Posters and Late Breaking Results, Workshops and Tutorials, Industry and Innovation Tracks, Practitioners, Doctoral Consortium and Blue Sky* (pp. 11–19). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-64312-5_2

Cooke, N. J., Cohen, M. C., Fazio, W. C., Inderberg, L. H., Johnson, C. J., Lematta, G. J., Peel, M., & Teo, A. (2024). From Teams to Teamness: Future Directions in the Science of Team Cognition. *Human Factors*, *66*(6), 1669–1680. https://doi.org/10.1177/00187208231162449

Corallo, A., Lazoi, M., & Secundo, G. (2012). Inter-organizational knowledge integration in Collaborative NPD projects: Evidence from the aerospace industry. *Knowledge Management Research & Practice*, *10*(4), 354–367. https://doi.org/10.1057/kmrp.2012.25

Corbin, J. M., & Strauss, A. (1990). Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative Sociology*, *13*(1), 3–21. https://doi.org/10.1007/BF00988593

Cowan, R. (2001). Expert systems: Aspects of and limitations to the codifiability of knowledge. *Research Policy*, *30*(9), 1355–1372. https://doi.org/10.1016/S0048-7333(01)00156-1

Creswell, J. W., & Clark, V. L. P. (2017). *Designing and conducting mixed methods research*. Sage publications. https://books.google.com/books?hl=zh-

CN&lr=&id=BXEzDwAAQBAJ&oi=fnd&pg=PP1&dq=Creswell+%26+Clark+(2017)&ots=UlzeNkisoF&sig=U_vgeSWRg0RvMoyYdgTznvCwX_Y

Cuskley, C., Woods, R., & Flaherty, M. (2024). The Limitations of Large Language Models for Understanding Human Language and Cognition. *Open Mind*, *8*, 1058–1083. https://doi.org/10.1162/opmi_a_00160

Danish, J. A., Enyedy, N., Saleh, A., & Humburg, M. (2020). Learning in embodied activity framework: A sociocultural framework for embodied cognition. *International Journal of Computer-Supported Collaborative Learning*, *15*(1), 49–87. https://doi.org/10.1007/s11412-020-09317-3

DeChurch, L. A., & Mesmer-Magnus, J. R. (2010). The cognitive underpinnings of effective teamwork: A meta-analysis. *Journal of Applied Psychology*, *95*(1), 32–53. https://doi.org/10.1037/a0017328

Dell'Era, C., Magistretti, S., Cautela, C., Verganti, R., & Zurlo, F. (2020). Four kinds of design thinking: From ideating to making, engaging, and criticizing. *Creativity and Innovation Management*, *29*(2), 324–344. https://doi.org/10.1111/caim.12353

Demetriadis, S., & Dimitriadis, Y. (2023). Conversational Agents and Language Models that Learn from Human Dialogues to Support Design Thinking. In C. Frasson, P. Mylonas, & C. Troussas (Eds.), *Augmented Intelligence and Intelligent Tutoring Systems* (pp. 691–700). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-32883-1_60

Eccles, D. W., & Arsal, G. (2017). The think aloud method: What is it and how do I use it? *Qualitative Research in Sport, Exercise and Health*, *9*(4), 514–531. https://doi.org/10.1080/2159676X.2017.1331501

Edwards, J. S., Shaw, D., & Collier, P. M. (2005). Knowledge management systems: Finding a way with technology. *Journal of Knowledge Management*, *9*(1), 113–125. https://doi.org/10.1108/13673270510583009

Ericsson, K. A. (2017). Protocol Analysis. In *A Companion to Cognitive Science* (pp. 425–432). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781405164535.ch33

Fan, I.-S., Li, G., Lagos-Hernandez, M., Bermell-Garci´a, P., & Twelves, M. (2008). *A Rule Level Knowledge Management System for Knowledge Based Engineering Applications*. 813–821. https://doi.org/10.1115/DETC2002/CIE-34501

Fauconnier, G. (1994). *Mental Spaces: Aspects of Meaning Construction in Natural Language*. Cambridge University Press.

Feng, K. J. K., Liao, Q. V., Xiao, Z., Wortman Vaughan, J., Zhang, A. X., & McDonald, D. W. (2025). Canvil: Designerly Adaptation for LLM-Powered User Experiences. *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3706598.3713139

Fiore, S. M., & Wiltshire, T. J. (2016). Technology as Teammate: Examining the Role of External Cognition in Support of Team Cognitive Processes. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01531

Galesic, M., Barkoczi, D., Berdahl, A. M., Biro, D., Carbone, G., Giannoccaro, I., Goldstone, R. L., Gonzalez, C., Kandler, A., Kao, A. B., Kendal, R., Kline, M., Lee, E., Massari, G. F., Mesoudi, A., Olsson, H., Pescetelli, N., Sloman, S. J., Smaldino, P. E., & Stein, D. L. (2023). Beyond collective intelligence: Collective adaptation. *Journal of The Royal Society Interface*, *20*(200), 20220736. https://doi.org/10.1098/rsif.2022.0736

Gao, C., Lan, X., Li, N., Yuan, Y., Ding, J., Zhou, Z., Xu, F., & Li, Y. (2024). Large language models empowered agent-based modeling and simulation: A survey and perspectives. *Humanities and Social Sciences Communications*, *11*(1), 1259. https://doi.org/10.1057/s41599-024-03611-3

Ge, S., Sun, Y., Cui, Y., & Wei, D. (2025). An Innovative Solution to Design Problems: Applying the Chain-of-Thought Technique to Integrate LLM-Based Agents With Concept Generation Methods. *IEEE Access*, *13*, 10499–10512. https://doi.org/10.1109/ACCESS.2024.3494054

Gibson, C. B. (2001). From knowledge accumulation to accommodation: Cycles of collective cognition in work groups. *Journal of Organizational Behavior*, *22*(2), 121–134. https://doi.org/10.1002/job.84

Gross, T. (2013). Supporting Effortless Coordination: 25 Years of Awareness Research. *Computer Supported Cooperative Work (CSCW)*, *22*(4), 425–474. https://doi.org/10.1007/s10606-013-9190-x

Hao, X., Demir, E., & Eyers, D. (2024). Exploring collaborative decision-making: A quasi-experimental study of human and Generative AI interaction. *Technology in Society*, *78*, 102662. https://doi.org/10.1016/j.techsoc.2024.102662

Heintz, C., & Scott-Phillips, T. (2023). Expression unleashed: The evolutionary and cognitive foundations of human communication. *Behavioral and Brain Sciences*, *46*, e1. https://doi.org/10.1017/S0140525X22000012

Herm, L.-V., Heinrich, K., Wanner, J., & Janiesch, C. (2023). Stop ordering machine learning algorithms by their explainability! A user-centered investigation of performance and explainability. *International Journal of Information Management*, *69*, 102538. https://doi.org/10.1016/j.ijinfomgt.2022.102538

Hirschberg, J., & Manning, C. D. (2015). Advances in natural language processing. *Science*, *349*(6245), 261–266. https://doi.org/10.1126/science.aaa8685

Ho, M.-T., & Vuong, Q.-H. (2025). Five premises to understand human–computer interactions as AI is changing the world. *AI & SOCIETY*, *40*(2), 1161–1162. https://doi.org/10.1007/s00146-024-01913-3

Hoff, K. A., & Bashir, M. (2015). Trust in Automation: Integrating Empirical Evidence on Factors That Influence Trust. *Human Factors*, *57*(3), 407–434. https://doi.org/10.1177/0018720814547570

Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*, *6*(2), 65–70.

Huang, C., Deng, Y., Lei, W., Lv, J., Chua, T.-S., & Huang, J. (2025). How to Enable Effective Cooperation Between Humans and NLP Models: A Survey of Principles, Formalizations, and Beyond. In W. Che, J. Nabende, E. Shutova, & M. T. Pilehvar (Eds.), *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 466–488). Association for Computational Linguistics. https://doi.org/10.18653/v1/2025.acl-long.22

Hurtienne, J. (2009). Cognition in HCI: An Ongoing Story. *Human Technology*, *5*(1), 12–28. https://doi.org/10.17011/ht/urn.20094141408

Iskra, M., Voigt, L., & Raab, M. (2024). Accounting for dynamic cognition–action interaction in decision-making tasks in sports: A scoping review. *Sport, Exercise, and Performance Psychology*, No Pagination Specified-No Pagination Specified. https://doi.org/10.1037/spy0000361

Järvelä, S., Kirschner, P. A., Hadwin, A., Järvenoja, H., Malmberg, J., Miller, M., & Laru, J. (2016). Socially shared regulation of learning in CSCL: Understanding and prompting individual- and group-level shared regulatory activities. *International Journal of Computer-Supported Collaborative Learning*, *11*(3), 263–280. https://doi.org/10.1007/s11412-016-9238-2

Jin, X., Dong, H., Evans, M., & Yao, A. (2024). Inspirational Stimuli to Support Creative Ideation for the Design of Artificial Intelligence-Powered Products. *Journal of Mechanical Design*, *146*(121402). https://doi.org/10.1115/1.4065696

Jin, Y., & Benami, O. (2010). Creative patterns and stimulation in conceptual design. *AI EDAM*, *24*(2), 191–209. https://doi.org/10.1017/S0890060410000053

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*. https://doi.org/10.1126/science.aaa8415

Kang, H. B., Lin, D. C.-E., Chen, Y.-Y., Hong, M. K., Martelaro, N., & Kittur, A. (2025). BioSpark: Beyond Analogical Inspiration to LLM-augmented Transfer. *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3706598.3714053

Kernan Freire, S., Wang, C., Foosherian, M., Wellsandt, S., Ruiz-Arenas, S., & Niforatos, E. (2024). Knowledge sharing in manufacturing using LLM-powered tools: User study and model benchmarking. *Frontiers in Artificial Intelligence*, *7*. https://doi.org/10.3389/frai.2024.1293084

Kim, J., & Ryu, H. (2014). A Design Thinking Rationality Framework: Framing and Solving Design Problems in Early Concept Generation. *Human–Computer Interaction*, *29*(5–6), 516–553. https://doi.org/10.1080/07370024.2014.896706

Kim, M. H., Kim, Y. S., Lee, H. S., & Park, J. A. (2007). An underlying cognitive aspect of design creativity: Limited Commitment Mode control strategy. *Design Studies*, *28*(6), 585–604. https://doi.org/10.1016/j.destud.2007.04.006

Kolasani, S. (2023). Optimizing Natural Language Processing, Large Language Models (LLMs) for Efficient Customer Service, and hyper-personalization to enable sustainable growth and revenue. *Transactions on Latest Trends in Artificial Intelligence*, *4*(4). https://ijsdcs.com/index.php/TLAI/article/view/476

Kozar, O. (2010). Towards Better Group Work: Seeing the Difference between Cooperation and Collaboration. *English Teaching Forum*, *48*(2), 16–23.

Kyriakopoulos, K., & De Ruyter, K. (2004). Knowledge Stocks and Information Flows in New Product Development. *Journal of Management Studies*, *41*(8), 1469–1498. https://doi.org/10.1111/j.1467-

6486.2004.00482.x

Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, *33*(1), 159–174. https://doi.org/10.2307/2529310

Lane, D., & Seery, N. (2011). *Examining the development of sketch thinking and behaviour* [Conference Proceedings]. 118th ASEE Annual Conference and Exposition. https://doi.org/10/1/Lane%252520and%252520Seery%2525202011%252520Examining_the_Development_of_Sketch_Thinking_and_Behaviour_ASEE_2011%2525281%252529.pdf

Lebiere, C., Blaha, L. M., Fallon, C. K., & Jefferson, B. (2021). Adaptive Cognitive Mechanisms to Maintain Calibrated Trust and Reliance in Automation. *Frontiers in Robotics and AI*, *8*. https://doi.org/10.3389/frobt.2021.652776

Leblebici-Başar, D., & Altarriba, J. (2013). The Role of Imagery and Emotion in the Translation of Concepts into Product Form. *The Design Journal*, *16*(3), 295–314. https://doi.org/10.2752/175630613X13660502571787

Lee, J. D., & See, K. A. (2004). Trust in Automation: Designing for Appropriate Reliance. *Human Factors*, *46*(1), 50–80. https://doi.org/10.1518/hfes.46.1.50_30392

Lee, J., Gu, N., & Williams, A. P. (2014). Parametric Design Strategies for the Generation of Creative Designs. *International Journal of Architectural Computing*, *12*(3), 263–282. https://doi.org/10.1260/1478-0771.12.3.263

Lee, J. H., & Ostwald, M. J. (2025). Enhancing online design collaboration: Revealing cognitive and linguistic fusion in remote teamwork. *International Journal of Design Creativity and Innovation*, *13*(3), 139–158. https://doi.org/10.1080/21650349.2025.2481254

Lee, J. H., Ostwald, M. J., & Gu, N. (2020). Collaborative Design: Team Cognition and Communication. In J. H. Lee, M. J. Ostwald, & N. Gu (Eds.), *Design Thinking: Creativity, Collaboration and Culture* (pp. 113–145). Springer International Publishing. https://doi.org/10.1007/978-3-030-56558-9_5

Li, X., Wang, S., Zeng, S., Wu, Y., & Yang, Y. (2024). A survey on LLM-based multi-agent systems: Workflow, infrastructure, and challenges. *Vicinagearth*, *1*(1), 9. https://doi.org/10.1007/s44336-024-00009-2

Liikkanen, L. A., & Perttula, M. (2009). Exploring problem decomposition in conceptual design among novice designers. *Design Studies*, *30*(1), 38–59. https://doi.org/10.1016/j.destud.2008.07.003

Liu, J., Yao, Y., An, P., & Wang, Q. (2024). PeerGPT: Probing the Roles of LLM-based Peer Agents as Team Moderators and Participants in Children's Collaborative Learning. *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, 1–6. https://doi.org/10.1145/3613905.3651008

Lloyd, P., & Scott, P. (1994). Discovering the design problem. *Design Studies*, *15*(2), 125–140. https://doi.org/10.1016/0142-694X(94)90020-5

Lobo, J. R. A., Szejka, A. L., & Canciglieri Junior, O. (2025). Towards a Cognitive New Product Development Framework Based on Collaborative Engineering and Digital Technologies. In D. N. Šormaz, B. Bidanda, O. Alhawari, & Z. Geng (Eds.), *Intelligent Production and Industry 5.0 with Human Touch, Resilience, and Circular Economy* (pp. 77–86). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-77723-3_7

Lombard, M., Snyder-Duch, J., & Bracken, C. C. (2002). Content Analysis in Mass Communication: Assessment and Reporting of Intercoder Reliability. *Human Communication Research*, *28*(4), 587–604. https://doi.org/10.1111/j.1468-2958.2002.tb00826.x

London, K., & Singh, V. (2013). Integrated construction supply chain design and delivery solutions. *Architectural Engineering and Design Management*, *9*(3), 135–157. https://doi.org/10.1080/17452007.2012.684451

Marzi, G., & Balzano, M. (2025). Artificial intelligence and the reconfiguration of NPD Teams: Adaptability and skill differentiation in sustainable product innovation. *Technovation*, *145*, 103254. https://doi.org/10.1016/j.technovation.2025.103254

Mathieu, J. E., Heffner, T. S., Goodwin, G. F., Salas, E., & Cannon-Bowers, J. A. (2000). The influence of shared mental models on team process and performance. *Journal of Applied Psychology*, *85*(2), 273–283. https://doi.org/10.1037/0021-9010.85.2.273

Murtaza, M., Cheng, C.-T., Fard, M., & Zeleznikow, J. (2024). Transforming Driver Education: A Comparative Analysis of LLM-Augmented Training and Conventional Instruction for Autonomous Vehicle Technologies. *International Journal of Artificial Intelligence in Education*. https://doi.org/10.1007/s40593-

024-00407-z

Nagaraj, V., Berente, N., Lyytinen, K., & Gaskin, J. (2020). Team Design Thinking, Product Innovativeness, and the Moderating Role of Problem Unfamiliarity. *Journal of Product Innovation Management*, *37*(4), 297–323. https://doi.org/10.1111/jpim.12528

Nahar, N., Zhou, S., Lewis, G., & Kästner, C. (2022). Collaboration challenges in building ML-enabled systems: Communication, documentation, engineering, and process. *Proceedings of the 44th International Conference on Software Engineering*, 413–425. https://doi.org/10.1145/3510003.3510209

Nejjar, M., Zacharias, L., Stiehle, F., & Weber, I. (n.d.). LLMs for science: Usage for code generation and data analysis. *Journal of Software: Evolution and Process*, *n/a*(n/a), e2723. https://doi.org/10.1002/smr.2723

Omidvar Tehrani, B., M, I., & Anubhai, A. (2024). Evaluating Human-AI Partnership for LLM-based Code Migration. *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, 1–8. https://doi.org/10.1145/3613905.3650896

Onnasch, L., Wickens, C. D., Li, H., & Manzey, D. (2014). Human Performance Consequences of Stages and Levels of Automation: An Integrated Meta-Analysis. *Human Factors*, *56*(3), 476–488. https://doi.org/10.1177/0018720813501549

Ostergaard, K. J., & Summers, J. D. (2008). *A Taxonomy for Collaborative Design*. 755–764. https://doi.org/10.1115/DETC2003/DAC-48781

Oygür, I. (2018). The machineries of user knowledge production. *Design Studies*, *54*, 23–49. https://doi.org/10.1016/j.destud.2017.10.002

Parasuraman, R., & Manzey, D. H. (2010). Complacency and Bias in Human Use of Automation: An Attentional Integration. *Human Factors*, *52*(3), 381–410. https://doi.org/10.1177/0018720810376055

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, *30*(3), 286–297. https://doi.org/10.1109/3468.844354

Patel, D., & Alismail, A. (2024). Relationship Between Cognitive Load Theory, Intrinsic Motivation and Emotions in Healthcare Professions Education: A Perspective on the Missing Link. *Advances in Medical*

Education and Practice, *15*, 57–62. https://doi.org/10.2147/AMEP.S441405

Pratt, W., Reddy, M. C., McDonald, D. W., Tarczy-Hornoch, P., & Gennari, J. H. (2004). Incorporating ideas from computer-supported cooperative work. *Journal of Biomedical Informatics*, *37*(2), 128–137. https://doi.org/10.1016/j.jbi.2004.04.001

Qin, J., van der Rhee, B., Venkataraman, V., & Ahmadi, T. (2021). The impact of IT infrastructure capability on NPD performance: The roles of market knowledge and innovation process formality. *Journal of Business Research*, *133*, 252–264. https://doi.org/10.1016/j.jbusres.2021.04.072

Qiu, Y., & Jin, Y. (2025). A Method for Synthesizing Ontology-Based Textual Design Datasets: Evaluating the Potential of Large Language Model in Domain-Specific Dataset Generation. *Journal of Mechanical Design*, *147*(4). Scopus. https://doi.org/10.1115/1.4067478

Razmerita, L., Kirchner, K., & Nabeth, T. (2014). Social Media in Organizations: Leveraging Personal and Collective Knowledge Processes. *Journal of Organizational Computing and Electronic Commerce*, *24*(1), 74–93. https://doi.org/10.1080/10919392.2014.866504

Riyadh, H. A., Khrais, L. T., Alfaiza, S. A., & Sultan, A. A. (2021). Association between mass collaboration and knowledge management: A case ofJordan companies. *International Journal of Organizational Analysis*, *31*(4), 973–987. https://doi.org/10.1108/IJOA-08-2021-2893

Salma, Z., Hijón-Neira, R., & Pizarro, C. (2025). Designing Co-Creative Systems: Five Paradoxes in Human–AI Collaboration. *Information*, *16*(10), 909. https://doi.org/10.3390/info16100909

Sarker, I. H. (2021). Data Science and Analytics: An Overview from Data-Driven Smart Computing, Decision-Making and Applications Perspective. *SN Computer Science*, *2*(5), 377. https://doi.org/10.1007/s42979-021-00765-8

Scharowski, N., Perrig, S. A. C., Svab, M., Opwis, K., & Brühlmann, F. (2023). Exploring the effects of human-centered AI explanations on trust and reliance. *Frontiers in Computer Science*, *5*. https://doi.org/10.3389/fcomp.2023.1151150

Shmueli, G., & Koppius, O. R. (2011). Predictive Analytics in Information Systems Research. *MIS Quarterly*, *35*(3), 553–572. https://doi.org/10.2307/23042796

Shteynberg, G., Hirsh, J. B., Wolf, W., Bargh, J. A., Boothby, E. J., Colman, A. M., Echterhoff, G., &

Rossignac-Milon, M. (2023). Theory of collective mind. *Trends in Cognitive Sciences*, *27*(11), 1019–1031. https://doi.org/10.1016/j.tics.2023.06.009

Shu-Hsien Liao. (2005). Expert system methodologies and applications—A decade review from 1995 to 2004. *Expert Systems with Applications*, *28*(1), 93–103. https://doi.org/10.1016/j.eswa.2004.08.003

Stauffer, L. A., & Ullman, D. G. (1991). Fundamental Processes of Mechanical Designers Based on Empirical Data. *Journal of Engineering Design*, *2*(2), 113–125. https://doi.org/10.1080/09544829108901675

Stempfle, J., & Badke-Schaub, P. (2002). Thinking in design teams—An analysis of team communication. *Design Studies*, *23*(5), 473–496. https://doi.org/10.1016/S0142-694X(02)00004-2

Sternberg, R. J. (1998). *Handbook of Creativity*. Cambridge University Press.

Suh, S., Chen, M., Min, B., Li, T. J.-J., & Xia, H. (2024). Luminate: Structured Generation and Exploration of Design Space with Large Language Models for Human-AI Co-Creation. *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3613904.3642400

Tan, J. (2010). Grounded theory in practice: Issues and discussion for new qualitative researchers. *Journal of Documentation*, *66*(1), 93–112. https://doi.org/10.1108/00220411011016380

Trovato, M., Belluomo, L., Bici, M., Prist, M., Campana, F., & Cicconi, P. (2025). Machine learning in design for additive manufacturing: A state-of-the-art discussion for a support tool in product design lifecycle. *The International Journal of Advanced Manufacturing Technology*, *137*(5), 2157–2180. https://doi.org/10.1007/s00170-025-15273-9

Tseng, S.-M. (2008). The effects of information technology on knowledge management systems. *Expert Systems with Applications*, *35*(1), 150–160. https://doi.org/10.1016/j.eswa.2007.06.011

Ugbebor, F. O., Adeteye, D. A., & Ugbebor, J. O. (2024). PREDICTIVE ANALYTICS MODELS FOR SMES TO FORECAST MARKET TRENDS, CUSTOMER BEHAVIOR, AND POTENTIAL BUSINESS RISKS. *Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (Online)*, *3*(3), 355–381. https://doi.org/10.60087/jklst.v3.n3.p355-381

Urquhart, C., Lehmann, H., & Myers, M. D. (2010). Putting the 'theory' back into grounded theory: Guidelines for grounded theory studies in information systems. *Information Systems Journal*, *20*(4), 357–

381. https://doi.org/10.1111/j.1365-2575.2009.00328.x

Vlasceanu, M., Dyckovsky, A. M., & Coman, A. (2024). A Network Approach to Investigate the Dynamics of Individual and Collective Beliefs: Advances and Applications of the BENDING Model. *Perspectives on Psychological Science*, *19*(2), 444–453. https://doi.org/10.1177/17456916231185776

Vu, M. D., Wang, H., Chen, J., Li, Z., Zhao, S., Xing, Z., & Chen, C. (2024). GPTVoiceTasker: Advancing Multi-step Mobile Task Efficiency Through Dynamic Interface Exploration and Learning. *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology*. https://doi.org/10.1145/3654777.3676356

Wang, J., Tigelaar, D. E. H., Zhou, T., & Admiraal, W. (2023). The effects of mobile technology usage on cognitive, affective, and behavioural learning outcomes in primary and secondary education: A systematic review with meta-analysis. *Journal of Computer Assisted Learning*, *39*(2), 301–328. https://doi.org/10.1111/jcal.12759

Wei, R., Li, K., & Lan, J. (2024). Improving Collaborative Learning Performance Based on LLM Virtual Assistant. *2024 13th International Conference on Educational and Information Technology (ICEIT)*, 1–6. https://doi.org/10.1109/ICEIT61397.2024.10540942

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, *3*(2), 159–177. https://doi.org/10.1080/14639220210123806

Wohllebe, A., & Lagodka, C. (2024). Can ChatGPT Replace Human Content Writers in SEO Copywriting? An Empirical Comparison for Performance Measurement of Generative AI in Search Engine Optimization in E-Commerce. In T. Bolz & G. Schuster (Eds.), *Generative Künstliche Intelligenz in Marketing und Sales: Innovative Unternehmenspraxis: Insights, Strategien und Impulse* (pp. 303–313). Springer Fachmedien. https://doi.org/10.1007/978-3-658-45132-5_21

Woo, B. M., Tan, E., Yuen, F. L., & Hamlin, J. K. (2023). Socially evaluative contexts facilitate mentalizing. *Trends in Cognitive Sciences*, *27*(1), 17–29. https://doi.org/10.1016/j.tics.2022.10.003

Woolley, A. W. (2025). Generative AI and collaboration: Opportunities for cultivating collective intelligence. *Journal of Organization Design*. https://doi.org/10.1007/s41469-025-00199-z

Xu, X. (Tone), Konnova, A., Gao, B., Peng, C., Vo, D., & Dow, S. P. (2025). Productive vs. Reflective:

How Different Ways of Integrating AI into Design Workflows Affect Cognition and Motivation. *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3706598.3713649

Yang, Y., Chen, Y., Feng, X., Sun, D., & Pang, S. (2024). Investigating the mechanisms of analytics-supported reflective assessment for fostering collective knowledge. *Journal of Computing in Higher Education*, *36*(1), 242–273. https://doi.org/10.1007/s12528-024-09398-1

Ye, Y., You, H., & Du, J. (2023). Improved Trust in Human-Robot Collaboration With ChatGPT. *IEEE Access*, *11*, 55748–55754. https://doi.org/10.1109/ACCESS.2023.3282111

Zahra, S. A., Neubaum, D. O., & Larrañeta, B. (2007). Knowledge sharing and technological capabilities: The moderating role of family involvement. *Journal of Business Research*, *60*(10), 1070–1079. https://doi.org/10.1016/j.jbusres.2006.12.014

Zamfirescu-Pereira, J. D., Jun, E., Terry, M., Yang, Q., & Hartmann, B. (2025). Beyond Code Generation: LLM-supported Exploration of the Program Design Space. *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3706598.3714154

Zhang, P., & Soergel, D. (2014). Towards a comprehensive model of the cognitive process and mechanisms of individual sensemaking. *Journal of the Association for Information Science and Technology*, *65*(9), 1733–1756. https://doi.org/10.1002/asi.23125

Zhang, Z., Peng, W., Chen, X., Cao, L., & Li, T. J.-J. (2025). LADICA: A Large Shared Display Interface for Generative AI Cognitive Assistance in Co-located Team Collaboration. *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3706598.3713289

Zhou, J., Li, R., Tang, J., Tang, T., Li, H., Cui, W., & Wu, Y. (2024). Understanding Nonlinear Collaboration between Human and AI Agents: A Co-design Framework for Creative Design. *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, 1–16. https://doi.org/10.1145/3613904.3642812

Figure legends

Figure 1. The 2I2A Model of Collective Cognition

Figure 2. Illustrative Example of 2I2A Coding Across Collaborative Spaces