ARTICLE

Check for updates

Inspiration booster or creative fixation? The dual mechanisms of LLMs in shaping individual creativity in tasks of different complexity

Xusen Cheng[1 ✉] & Lulu Zhang[1 ✉]

The emergence of large language models (LLMs) presents opportunities for stimulating unlimited creative potential. However, how LLMs influence individual creativity remains unclear. Therefore, this paper examines the dual-opposing mechanisms through which LLMs influence individual creativity. In Experiment 1, each participant collaborated with a human partner or a general, unconstrained LLM partner to complete creative tasks. The results showed that compared to collaborating with the human partner, collaborating with the LLM partner significantly improved individual creativity in simple tasks, attributable to inspiration stimulation. However, in complex tasks, collaborating with the LLM partner led to a decrease in creativity, attributable to creative fixation. To mitigate this impact, in Experiment 2, participants were instructed to collaborate with batch-responsive LLM or constrained-responsive LLM to complete creative tasks. We found that constraining the output of LLMs effectively mitigated the creative fixation they induce in complex tasks, thereby enhancing creative performance. However, this constraint may weaken the positive effects of inspiration stimulation in simple tasks. These findings provide insights for the differentiated application of LLMs in creative tasks.

[1] School of Information, Renmin University of China, Beijing, China. ✉email: xusen.cheng@ruc.edu.cn; zhanglulu0617@163.com

Introduction

Since the launch of ChatGPT, large language models (LLMs) have been extensively utilized to tackle creative tasks across various domains such as education, business, engineering, design and literary creation (Lee and Chung 2024; Anderson et al. 2024; Beguš 2024). ChatGPT has not only excelled in a wide range of creative tasks, such as coming up with research ideas and generating marketing slogans (Conroy 2024; Cillo and Rubera 2024), but also made significant contributions to enhancing human creative performances. For example, management scholars are increasingly examining how organizations can utilize LLM tools to bolster employee creativity (Jia et al. 2024). Education researchers believe that LLM-supported tools can encourage students' creativity and problem-solving skills (Extance 2023). Creativity is a vital force driving transformation and development across numerous industries (Janger et al. 2017; Boussioux et al. 2024). Therefore, one of our objectives is to promote the use of LLMs to better assist individuals or teams in proposing more creative ideas and solutions in creative tasks, thereby infusing innovative energy into sectors of society (Epstein et al. 2023).

In addition to using LLMs as tools to assist in solving creative tasks, the collaboration between humans and LLMs in creative tasks is believed to unlock unprecedented creative potential (Boussioux et al. 2024; Rafner et al. 2023). Hybrid intelligence may surpass the mode of human or AI working alone, achieving enhancement in creativity (Rafner et al. 2023). It is important to note that LLMs, when serving as tools or collaborators, are different (Liu et al. 2024). The former emphasizes using LLMs as answer-producing tools to improve the efficiency and quality of creative output, while the latter emphasizes the cultivation of creative thinking ability based on collaborative processes (Liu et al. 2024; Eapen et al. 2023). In fact, LLM tools like ChatGPT indeed assist in completing creative tasks more effectively, but their impact on cultivating long-term creative thinking abilities is controversial (Liu et al. 2024). The main bottleneck lies in the coordination between LLMs, humans, and tasks, as well as in finding optimal interaction mode (Rafner et al. 2023; McGill and Klobas 2009). Therefore, this study aims to explore a balanced collaboration pattern between leveraging LLMs to enhance the efficiency of real-time creative solution generation while ensuring the development of human creative thinking abilities.

Currently, empirical research on the impact of generative AI tools such as LLMs on creativity in creative tasks is still in its early stages (Lee and Chung 2024). Some scholars believe that LLMs facilitate the innovative combination of existing knowledge bases and distantly related concepts, ultimately improving human performance in various daily creative tasks (Jia et al. 2024; Bouschery et al. 2024). However, other studies suggest that LLMs may stifle collective diversity of ideas and creative outputs, lead to content homogenization, and ultimately harm the development of creativity and scientific progress (Doshi and Hauser 2024; Nakadai et al. 2023). Overall, the existing opinions regarding the impact of LLMs on creativity are diverse but relatively scattered (Lee and Chung 2024), and exploration of the negative effects at the individual level is particularly limited. The complex mechanisms by which LLMs affect creativity have not been systematically explored. Therefore, we explored the following question: How do LLMs influence human creativity through dual oppositional pathways? And what is the boundary condition of the pathway?

Creativity is considered to be the innovative combination of seemingly unrelated knowledge or concepts to form new ideas (Lee and Chung 2024). Previous studies have shown that human-human collaboration, represented by brainstorming, is an important way to inspire inspiration and creativity (Osborn 1953;

Hwang and Won 2021; Dennis and Valacich 1993), however, the collaboration may result in additional process loss, such as distraction effect, attentional production blocking, striving for originality, cognitive complexity, cognitive dispersion (Pinsonneault et al. 1999; DeSanctis and Gallupe 1987; Karau and Williams 1993; Stephen et al. 2016). LLMs provide an opportunity to mitigate these negative impacts (Hwang and Won 2021; Dennis and Valacich 1993). Furthermore, considering the controllability and excellent performance of LLMs, they have the potential to become excellent collaborators (Eapen et al. 2023; Bouschery et al. 2024). However, it is currently unclear, how psychological and behavioral characteristics observed in human-human collaboration transfer to human-LLM setting. Therefore, this paper comprehensively reconsiders the positive and negative impacts of LLMs, which may revolutionize our understanding of creativity (Epstein et al. 2023; Rafner et al. 2023). Specifically, this study constructs a dual oppositional model of how LLMs influence individual creativity, and introduces task complexity as a moderating variable to explore how LLMs shape creativity in different task contexts.

The first highlight of this article is that we shift our focus from long-term collective negative impacts to short-term individual level, exploring how creative fixation undermines individual creative performance in single interaction and seeking ways to alleviate this negative impact. Recent studies have demonstrated that LLMs can lead to idea homogenization at the collective level, raising concerns about the negative impacts of their long-term use, such as LLM potentially eroding our brains and thinking (Anderson et al. 2024). To curb this long-term impact, it is crucial to suppress the accumulation of potential negative effects of LLM in each short-term interaction, which is exactly what our research has done. Secondly, we have considered the relative magnitude of the negative impact and the positive impact under different task complexity conditions, innovatively revealing the opposing pathways and boundary conditions through which LLMs influence creativity based on cognitive load theory. Finally, our results demonstrate the necessity of differentially deploying LLM tools based on characteristics of creative tasks. This has far-reaching implications for the development and improvement of AI-based creativity support tools in various domains.

Theoretical background and hypothesis development

Cognitive load theory. To systematically investigate how LLMs impact human creativity, we applied cognitive load theory and its extension, collaborative cognitive load theory. According to the cognitive load theory, new information is processed in working memory, a process crucial for creativity. During this phase, two types of cognitive load sources are imposed on working memory (Sweller 2010). The intrinsic load comes from the inherent complexity of the task, determined by the number of new information elements in the task and their interrelationships (Kirschner et al. 2018). In addition, extraneous load arises from interacting elements unrelated to the inherent complexity (Kirschner et al. 2018). Cognitive overload occurs when these loads exceed working memory capacity (Kirschner et al. 2018). Furthermore, collaborative cognitive load theory examines how cognitive loads change during collaboration (Janssen and Kirschner 2020. On the one hand, collaboration can distribute the intrinsic load among group members' working memories, freeing up individual cognitive resources for deeper creative thinking (Janssen and Kirschner 2020); On the other hand, low-quality communication and distractions, etc., may introduce additional collaboration costs, leading to an increase in extraneous cognitive load (Potter and Balthazard 2004; Elen and Clark 2006).

These theories offer insights into how LLMs impact human creativity by allocating internal load and increasing external load. As collaborators, LLMs generate numerous ideas and information, boosting the likelihood of inspiring novel thoughts and distant associations, enhancing creativity (Doshi and Hauser 2024; Hofstetter et al. 2021). They undeniably serve as excellent internal load distributors or inspiration catalysts in human-LLM collaborations (Doshi and Hauser 2024; Janssen and Kirschner 2020). However, their remarkable capabilities of LLM may challenge individual thinking. When individuals are required to process an excessive number of interacting elements, the cognitive resources demanded may exceed individual working memory capacity, resulting in cognitive overload and impeding learning mechanisms (Janssen and Kirschner 2020). Additionally, we consider task complexity as a crucial factor affecting collaboration effectiveness, employing it as a boundary condition to explore different influencing mechanisms (Janssen and Kirschner 2020).

Dual pathways. Existing research shows that, on the one hand, LLMs excel at offering inspirational stimulation, thereby promoting associations and creative thinking (Doshi and Hauser 2024). Conversely, LLMs may also impose additional external cognitive load, a phenomenon we term creative fixation (Crilly 2019). During the early stages of ideation, LLMs present users with numerous highly structured, informative, and logical ideas (Urban et al. 2024). This deluge of information may result in individuals' limited working memory resources being occupied by a large number of existing interacting elements, leading to cognitive overload (Janssen and Kirschner 2020). At this point, lacking cognitive resources for additional thinking, individuals may consider the answers provided by LLMs are good enough or valuable without carefully judging (Anderson et al. 2024). Therefore, people spend more time processing ideas already proposed by LLMs, such as expansion and reorganization, rather than proposing new ideas. In this process, people unconsciously follow the logical framework of LLMs, which prevents individuals from escaping established notions and devising innovative solutions, ultimately resulting in creative fixation (Hofstetter et al. 2021).

Furthermore, in addition to external information shocks from LLMs, the inherent complexity of the task can also lead to an increase in interacting elements, thus, a complex task may accelerate the formation of the aforementioned cognitive overload and creative fixation (Janssen and Kirschner 2020). Moreover, in complex tasks, individuals' ability to come up with ideas or solutions within a certain time limit diminishes, making them more reliant on ready-made answers from LLMs rather than putting in more cognitive effort to think independently, thereby exacerbating creative fixation (Salimzadeh et al. 2024). Thus, the negative effects of LLMs outweigh the positive effects, reducing creativity. Conversely, in simple tasks, the generation of ideas becomes easier for individuals, prompting them to actively propose different ideas and improve the diversity of the idea set (Anderson et al. 2024). Therefore, the positive effects of LLMs outweigh the negative effects, and creativity is improved. Based on these analyses, we hypothesized that in complex tasks, collaborating with an LLM partner (vs a human partner) reduces creativity, and this negative impact is mediated by creative fixation; In simple tasks, collaborating with an LLM partner (vs a human partner) enhances creativity, and this positive impact is mediated by inspirational stimulation.

LLM response type. As one of the two opposing processes that underlie the impact of LLMs, creative fixation may harm the

development of human creativity. To alleviate this negative impact, according to the cognitive load theory, the core lies in reducing the interacting elements in collaboration (Sweller 2010). Specifically, appropriate constraints on the output of LLMs to align with the human information processing pace in specific tasks, thereby preventing an overwhelming surge of information (Hofstetter et al. 2021). Based on the previous hypothesis, LLMs tend to exacerbate creative fixation in tasks of high complexity. Thus, constraining the output of the LLM partner can prevent individuals from being overwhelmed by too much information, which allows individuals to retain ample idle cognitive resources within their limited working memory for creative thinking, thereby reducing creative fixation and augmenting creativity (Sweller 2010) (Fig. 1). Conversely, in low complexity tasks, constraining the output of the LLM partner may actually suppress creativity. Without exceeding the limit of working memory, greater exposure to external ideas stimulates divergent thinking and associative processes (Gallupe et al. 1992). Notably, simple tasks inherently contain fewer interacting elements, drastically minimizing the risk of exceeding working memory limits (Kirschner et al. 2011). Hence, constraining the output of the LLM partner in such tasks actually reduces exposure to external ideas, potentially reducing inspirational stimulation (Hofstetter et al. 2021). Consequently, we hypothesized that in complex tasks, collaborating with constrained-responsive (vs batch-responsive) LLM increases creativity, and this positive impact is mediated by creative fixation. In simple tasks, collaborating with constrained-responsive (vs batch-responsive) LLM reduces creativity, and this negative impact is mediated by inspirational stimulation.

Materials and methods

Study materials. This experiment was conducted online using our self-developed AI collaboration system. Each participant logged in using a pre-assigned username and entered a room with a specific task theme by inputting the corresponding room ID (see Supplementary Figs. S1 and 2). The system included functions such as task introduction, timer, chat box, and text box for submitting ideas (see Supplementary Fig. S3). The submitted ideas were displayed in a specific area and can be edited and deleted (see Supplementary Fig. S4). In addition, we integrated ERNIE Bot, the knowledge-enhanced large language model developed by Baidu, as an LLM partner. To meet the needs of current research, we configured LLM partners to follow specific scripts when proposing ideas. Specifically, we randomly selected 50 different ideas from the ideas proposed by ERNIE Bot in advance as the LLM idea pool (see Supplementary Tables S1 and 2). The ideas in the LLM idea pool for both simple and complex tasks were rated as having moderate creativity ($M_{complex} = 2.2, \text{s.d.} = 0.404; M_{easy} = 2.16, \text{s.d.} = 0.370$). In order to simulate the process of using LLMs to assist in generating ideas in reality, we configured LLM partners as to randomly select five ideas from the idea pool as responses each time. A time interval of 100 s was imposed between consecutive responses to ensure adequate reading time for the participants. Within the 10-min task duration, the LLM partners responded six times in total. In Experiment 2, we adjusted the configuration of the LLM partners to include either two or ten different ideas in each response. In both cases, we ensured that participants received a total of 30 ideas from their LLM partners. Under the condition of generating two ideas per response, the time interval between consecutive responses was set to 40 s (a total of 15 responses), and under the condition of generating 10 ideas per response, the time interval was extended to 200 s (a total of 3 responses).

When participants were assigned to collaborate with a human partner, LLM was excluded from the room. In order to eliminate
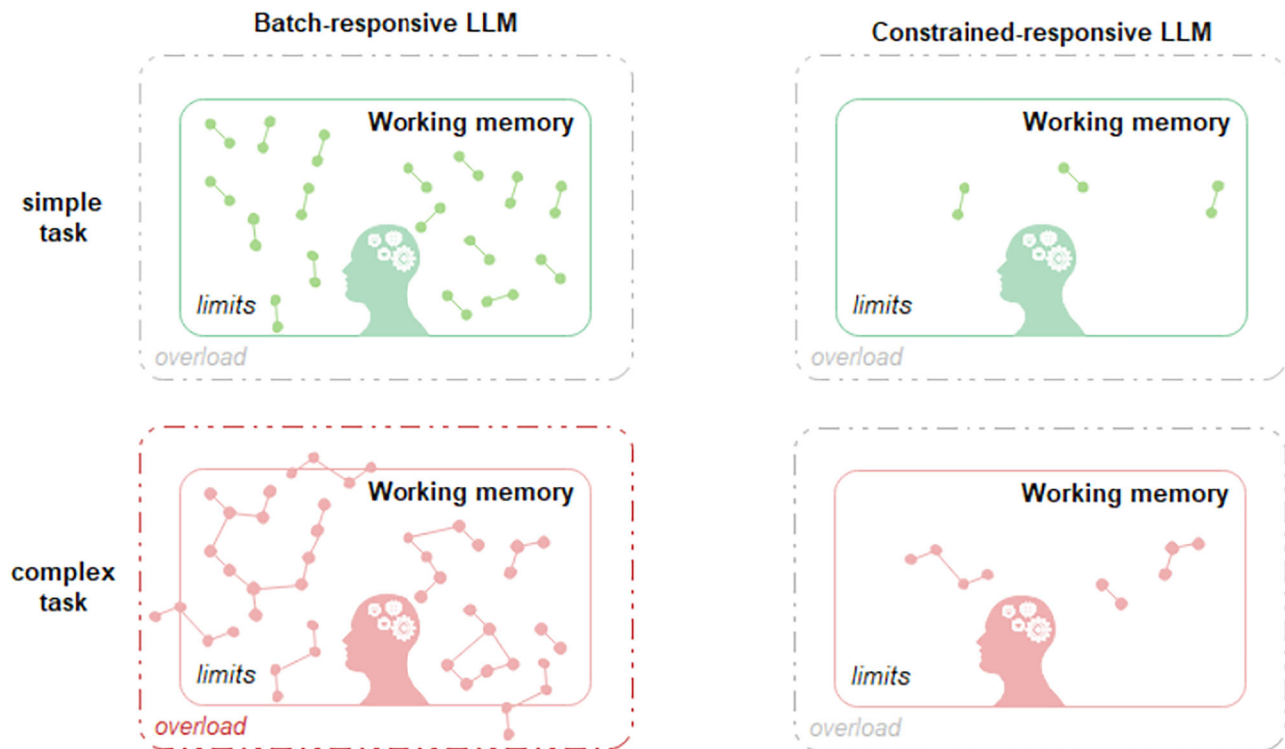
Fig. 1 A visual summary of research logic. The dots and lines represent interacting elements and their relationships. In simple tasks, there are fewer interacting elements that need to be processed simultaneously. Therefore, regardless of whether the number of ideas output by LLM is constrained, it is unlikely to exceed the limit of working memory. However, in complex tasks, there are many interacting elements that need to be processed. Batch-responsive LLM generates a large number of ideas in a short period of time, further increasing the number of interacting elements, which may exceed the limits of working memory and have negative effects. In this case, constraining the output of LLM will greatly reduce the risk of overload.

individual differences among human partners, participants interacted with the same research assistant. The research assistant followed specific scripts in the process of proposing ideas. The research assistant would first greet the participants and invite them to begin the task. During the 10-min task period, the research assistant would copy ideas from a human idea pool and paste them into the chat. The order of presenting the ideas would depend on their interaction and conversation with the participants. The research assistant could make appropriate modifications or expansions to the ideas and provide suitable responses to what the participants say in the chat box. The ideas in the human partner's idea pool originated from the responses of eight people in the researcher network during two creative tasks. After removing redundant ideas, 50 were randomly selected from their answers and stored (see Supplementary Tables S3 and 4). The ideas in the human idea pool for both simple and complex tasks are rated as having moderate creativity ($M_{complex} = 2.18$, s.d. = 0.388; $M_{easy} = 2.2$, s.d. = 0.404).

Furthermore, there was no significant difference in the creativity scores between the ideas in the LLM idea pool and the human idea pool, regardless of whether the tasks are simple ($M_{human} = 2.200$, s.d. = 0.404 vs $M_{LLM} = 2.160$, s.d. = 0.370; $t_{98} = 0.516$; $P = 0.607$; $d = 0.103$; 95% CI, (−0.1138, 0.1938)) or complex ($M_{human} = 2.180$, s.d. = 0.388 vs $M_{LLM} = 2.200$, s.d. = 0.404; $t_{98} = -0.252$; $P = 0.801$; $d = 0.050$; 95% CI, (−0.1772, 0.1372)).

In addition, we referred to creativity literature to manipulate task complexity. Specifically, the simple task is alternative use testing (come up with creative uses for the item), which contains fewer interacting elements. The complex task involves coming up with innovative features and proposing solutions for enterprise products, which include more interacting elements (Conroy 2024).

Procedure
Experiment 1
Participants and procedures: We recruited 204 participants in return for monetary compensation, requiring them to have experience using LLM tools. The average age of the participants was 27.73-years-old. 45.1% of the participants were male and 54.9% were female. Among them, 91.18% of the participants had a bachelor's degree or above. The demographic data of gender, age, education, and experience in using generative AI tools under different conditions can be seen in Supplementary Table S5.

This experiment was conducted online using our self-developed AI collaboration system. Participants were randomly assigned to one condition in the two (task complexity: low vs high) by two (partner: human vs LLM) between-participants design. Participants were first required to disclose their demographics and then log into our collaboration system to begin completing creative tasks with a human or a LLM partner (see Supplementary Section S5 for experimental procedure details). The simple task required participants to propose creative uses for the tire as an item. The complex task required participants to propose innovative feature points and solutions for the product (VR glasses) developed by the enterprise. During the task, participants obtained stimulation by reading ideas proposed by their human or LLM partner. Participants had 10 min to generate as many ideas as possible. After completing the task, the participants completed the manipulation check. Additionally, we measured participants' self-reported creativity of the generated ideas.

Measurements: For the measurement of creativity, this study simultaneously assesses self-reported creativity and objective creativity scores, aiming to fully capture the impact of LLMs on

individual creativity from the two dimensions of subjective perception and objective outcomes. Self-reported creativity was measured using a scale (e.g., "How do you rate your creative performance in this task?"). For objective creativity scores, two trained, independent external experts, who were blind to participants' demographic data, research hypotheses, and experimental design, coded the creativity. The two experts were asked to independently rate each idea based on its overall novelty and usefulness, using a 5-point rating scale (1 = extremely low, 5 = extremely high). The interrater agreement was high (ICC = 0.847), so we used their average score as the overall creativity score for each participant's ideas.

We referred to previous research to measure creative fixation (Lu et al. 2017; Shin et al. 2020). Two experts were asked to independently rate each idea based on its overall similarity to the ideas already generated by LLM. The expert was instructed to rate each participant's idea using a 3-point scale: 0 = the idea does not copy any keywords or ideas; 1 = the idea copies a keyword or idea, but expands or modifies it; and 2 = the idea nearly copies an existing one (Shin et al. 2020). The interrater agreement across the two raters was high (ICC = 0.885), therefore, we adopted their average score as the overall creative fixation score for each participant's ideas. Inspirational stimulation was measured using five items from Bottger et al. (e.g., "My imagination was stimulated," α = 0.968) (Böttger et al. 2017). All of these questions were measured on a 5-point scale.

Experiment 2
Participants and procedures: We recruited 204 participants in return for monetary compensation, requiring them to have experience using LLM tools. The average age of the participants was 29-years-old. 41.2% of the participants were male and 58.8% were female. Among them, 94.12% of the participants had a bachelor's degree or above. The demographic data of gender, age, education, and experience in using generative AI tools under different conditions can be seen in Supplementary Table S6.

This experiment was conducted online using our self-developed AI collaboration system. Participants were randomly assigned to one condition in the two (task complexity: low vs high) by two (LLM response type: batch-responsive vs constrained-responsive) between-participants design. Similarly, participants were first required to disclose their demographics, and then collaborated with a LLM partner in the collaborative system to complete creative task (batch-responsive LLM: each response contained ten different ideas vs constrained-responsive LLM: each response contained two different ideas) (Hofstetter et al. 2021). The simple task required participants to propose creative uses for the tire as an item. The complex task required participants to propose innovative feature points and solutions for the product (VR glasses) developed by the enterprise (see Supplementary Section S5 for experimental procedure details). During the task, participants obtained stimulation by reading ideas proposed by the LLM partner. Participants had 10 min to generate as many ideas as possible. After completing the task, the participants completed the manipulation check. Additionally, we measured participants' self-reported creativity of the generated ideas.

Measurements: Two well-trained experts independently coded creativity using the same scales used in Experiment 1. The interrater agreement across the two raters was high (ICC = 0.841), therefore we adopted their average score as the overall creativity score of each participant's ideas. Similarly, two experts independently coded creative fixation using the same scales used in Experiment 1. The interrater agreement across the two raters was high (ICC = 0.900), therefore we adopted their

average score as the overall creative fixation score of each participant's ideas. Inspirational stimulation was measured using five items from Bottger et al. (e.g., "My imagination was stimulated," α = 0.885) (Böttger et al. 2017). All of these questions were measured on a 5-point scale.

Results
Experiment 1
Manipulation checks. The results confirmed successful manipulations. Under complex task conditions, the task complexity was considered higher, while under simple task conditions, the task complexity was considered lower ($M_{complex} = 3.30$, s.d. $= 0.888$ vs $M_{easy} = 2.58$, s.d. $= 0.826$; $t_{202} = -6.044$; $P < 0.001$; $d = -0.84$; 95% CI, $(-0.962, -0.489)$).

Self-reported creativity. A one-way ANOVA revealed a significant main effect of the partner type ($F_{1,202} = 16.489$, $P < 0.001$, $\eta_p^2 = 0.075$). Participants collaborating with an LLM partner (compared to a human partner) rated their proposed ideas as less creative ($M_{human} = 3.880$, s.d. $= 0.891$ vs $M_{LLM} = 3.404$, s.d. $= 0.782$; $t_{202} = 4.061$; $P < 0.001$; $d = 0.569$; 95% CI, $(0.2449, 0.7074)$).

To explore the reasons behind this gap, we further examined the objective creativity scores (rated by experts) between the two groups using a one-way ANOVA. The results revealed a significant main effect of the partner type ($F_{1,202} = 4.638$, $P = 0.032$, $\eta_p^2 = 0.022$), the objective creativity score of the AI collaboration group was significantly higher than that of the human collaboration group ($M_{human} = 2.311$, s.d. $= 0.469$ vs $M_{LLM} = 2.45$, s.d. $= 0.453$; $t_{202} = -2.154$; $P = 0.032$; $d = -0.301$; 95% CI, $(-0.266, -0.012)$). This indicates that the actual creativity did not decline when collaborating with AI; instead, it even showed a slight increase.

We then calculated the gap as the difference between the self-reported creativity (subjective score) and the objective creativity score. An independent-samples t-test on this gap revealed a significant difference between groups ($M_{human} = 1.569$, s.d. $= 0.899$ vs $M_{LLM} = 0.954$, s.d. $= 0.868$; $t_{202} = 4.972$; $P < 0.001$; $d = -0.696$; 95% CI, $(0.371, 0.859)$: the human collaboration group showed a larger gap, indicating that participants working with human partners significantly overestimated their creativity relative to expert ratings. In contrast, the AI collaboration group exhibited a smaller gap.

Creativity. Our 2 (partner type) × 2 (task complexity) ANOVA analysis showed that partner type had a significant effect $\left(F_{1,200} = 4.901, P = 0.028, \eta_p^2 = 0.024\right)$, but the task complexity had no significant effect on creativity ($F_{1,200} = 0.232$, $P = 0.631$, $\eta_p^2 = 0.001$). As shown in Fig. 2, we also replicated a significant interaction effect between task complexity and partner type ($F_{1,200} = 13.272$, $P < 0.001$, $\eta_p^2 = 0.062$). In complex tasks, participants who collaborate with an LLM partner were rated as less creative in their ideas compared to those who collaborate with a human partner ($M_{LLM} = 2.351$, s.d. $= 0.62$ vs $M_{human} = 2.441$, s.d. $= 0.63$; $F_{1,200} = 1.021$; $P = 0.028$; $\eta_p^2 = 0.005$). In simple tasks, participants collaborating with an LLM partner were rated as more creative in their ideas compared to collaborating with a human partner ($M_{LLM} = 2.549$, s.d. $= 0.62$ vs $M_{human} = 2.182$, s.d. $= 0.63$; $F_{1,200} = 17.152$; $P < 0.001$; $\eta_p^2 = 0.079$).

Moderated mediation. To test whether creative fixation and inspirational stimulation mediate the role of partner type on participants' creativity, we ran a moderated parallel mediation

5

model with 10,000 bootstrapped estimates (Hayes 2015; Preacher and Hayes 2008; Preacher et al. 2007) (Fig. 3; for regressions, see Supplementary Table S7). Firstly, we found that collaborating with an LLM partner (compared to collaborating with a human partner) significantly increased creativity fixation in complex tasks ($\beta = 1.449$, SE $= 0.154$, $P < 0.001$). Creativity fixation was negatively correlated with creativity scores ($\beta = -0.222$, SE $= 0.072$, $P = 0.002$). Under complex task conditions, the conditional indirect effect was negatively significant ($\omega_{M1} = -0.322$, 95% CI, $(-0.536, -0.120)$), but not significant under simple task conditions ($\omega_{M1} = -0.003$, 95% CI, $(-0.044, 0.045)$); In addition, the index of moderated mediation was significant and negative ($\omega_{M1} = -0.319$, 95% CI, $(-0.551, -0.118)$). Inspirational stimulation showed a different pattern. We found that collaborating with an LLM partner (compared to collaborating with a human partner) positively influences cognitive stimulation ($\beta = 1.661$, SE $= 0.117$, $P < 0.001$). Inspiration stimulation was positively related to the creativity score ($\beta = 0.401$, SE $= 0.114$, $P = 0.001$). The indirect

effects were positively significant in both the complex task conditions ($\omega_{M2} = 0.629$, 95% CI, $(0.266, 1.052)$) and the simple task conditions ($\omega_{M2} = 0.665$, 95% CI, $(0.271, 1.133)$), while the index of moderated mediation was not significant ($\omega_{M2} = -0.036$, 95% CI, $(-0.202, 0.087)$).

Experiment 1 demonstrates that, compared to collaborating with a human partner, collaborating with an LLM partner can enhance humans' creative performance in simple tasks. This positive effect is attributed to the inspirational stimulation provided by LLM. However, in complex tasks, the creative fixation effect induced by LLM intensifies, leading to a reversal of the positive effect. To mitigate the negative impact of creative fixation, in Experiment 2, we introduce two types of LLMs with different response patterns as partners.

Experiment 2

Manipulation checks. The results confirmed successful manipulations. Under complex task conditions, the task complexity was considered higher, while under simple task conditions, the task complexity was considered lower ($M_{complex} = 3.206$, s.d. $= 0.905$ vs $M_{easy} = 2.324$, s.d. $= 0.798$; $t_{202} = -7.387$; $P < 0.001$; d $= 1.039$; 95% CI, $(-1.118, -0.647)$). Under the constrained-responsive LLM condition, participants perceived that the number of ideas contained in each response of the LLM is significantly less than that under the batch-responsive LLM condition ($M_{batch} = 4.040$, s.d. $= 1.024$ vs $M_{constrained} = 3.058$, s.d. $= 1.105$; $t_{202} = 6.580$; $P < 0.001$; d $= 0.921$; 95% CI, $(0.688, 1.278)$).

Self-reported creativity. A one-way ANOVA revealed that the main effect of the response type is not significant ($F_{1,202} = 0.226$, $P = 0.635$, $\eta_p^2 = 0.001$). There was no significant difference in the self-reported creativity of participants collaborating with batch-responsive LLM and constraint-responsive LLM ($M_{batch} = 3.303$, s.d. $= 0.792$ vs $M_{constrained} = 3.279$, s.d. $= 0.743$; $t_{202} = 0.476$; $P = 0.635$; d $= 0.067$; 95% CI, $(-0.1608, 0.2631)$).

Creativity. Our 2 (LLM response type) × 2 (task complexity) ANOVA analysis showed that LLM response type ($F_{1,200} = 0.811$, $P = 0.369$, $\eta_p^2 = 0.004$) and task complexity ($F_{1,200} = 0.049$, $P = 0.826$, $\eta_p^2 = 0.001$) both had no significant effect on creativity. As shown in Fig. 4, we also replicated a significant interaction effect between LLM response type and task complexity ($F_{1,200} = 49.984$, $P < 0.001$, $\eta_p^2 = 0.200$). In complex
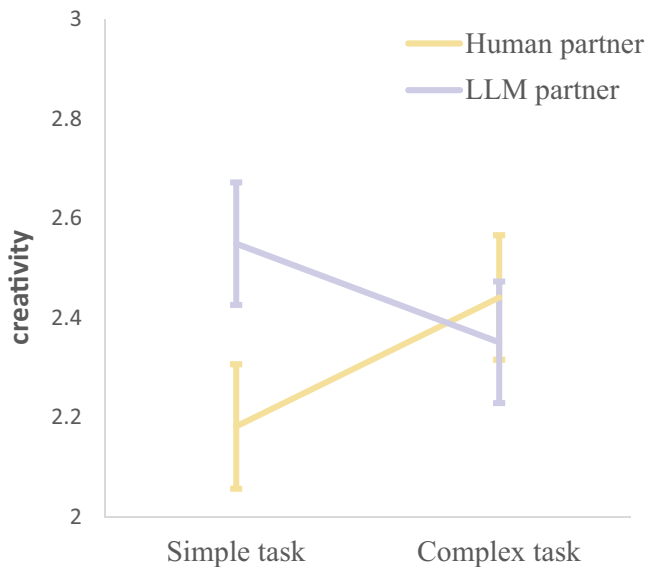

Fig. 2 The interaction effect between partner type and task complexity. In complex tasks, collaborating with an LLM partner (vs a human partner) reduces creativity; whereas in simple tasks, creativity is instead enhanced (study 1; $N = 204$). Error bars represent s.e.m.
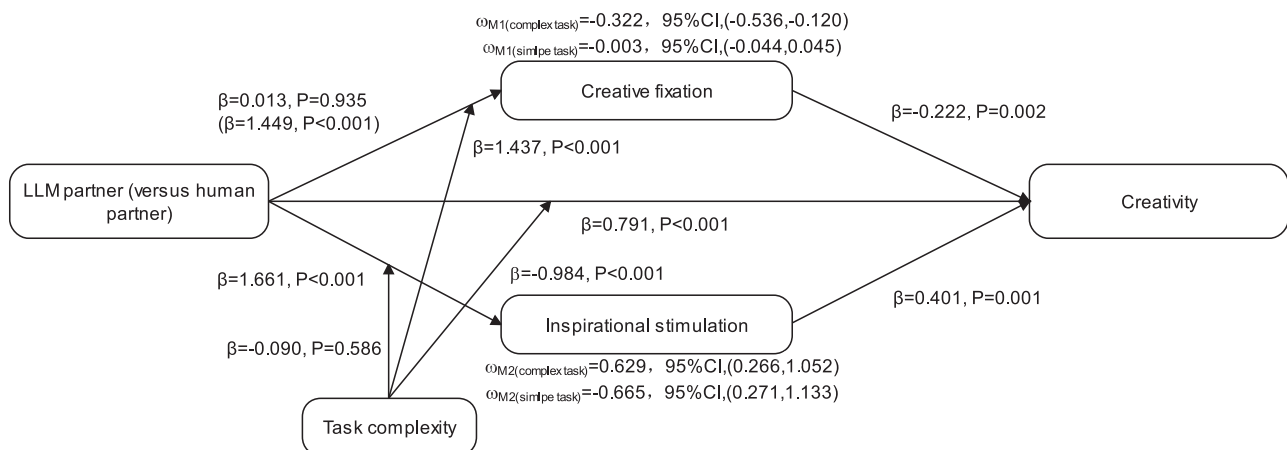

Fig. 3 Moderated mediation. A moderated parallel mediation model demonstrates that LLM influences human creativity through dual, opposing processes (inspiration, stimulation, and creative fixation). Numbers in parentheses are effects under a complex task. Standard errors are bootstrapped with 10,000 replications.

tasks, participants who collaborate with batch-responsive LLM were rated as less creative in their ideas compared to those who collaborate with constrained-responsive LLM ($M_{batch} = 2.139$, s.d. $= 0.048$ vs $M_{constrained} = 2.520$, s.d. $= 0.047$; $F_{1,200} = 31.765$; $P < 0.001$; $\eta_p^2 = 0.137$. In simple tasks, participants who collaborate with batch-responsive LLM were rated as more creative in their ideas compared to those who collaborate with constrained-responsive LLM ($M_{batch} = 2.466$, s.d. $= 0.048$ vs $M_{constrained} = 2.171$, s.d. $= 0.047$; $F_{1,200} = 19.031$; $P < 0.001$; $\eta_p^2 = 0.087$).

Moderated mediation. To test whether creative fixation and inspirational stimulation mediate the role of LLM response type on participants' creativity, we ran a moderated parallel mediation model with 10,000 bootstrapped estimates (Hayes 2015; Preacher and Hayes 2008; Preacher et al. 2007) (Fig. 5; for regressions, see Supplementary Table S8). Firstly, we found that collaborating

Fig. 4 The interaction effect between LLM response type and task complexity. In complex tasks, collaborating with constrained-responsive LLM (vs batch-responsive LLM) enhances creativity; whereas in simple tasks, creativity is instead reduced (study 2; $N = 204$). Error bars represent s.e.m.

with constrained-responsive LLM (compared to collaborating with batch-responsive LLM) significantly decreased creativity fixation in complex tasks ($\beta = -1.433$, SE $= 0.159$, $P < 0.001$). Creativity fixation was negatively correlated with creativity scores ($\beta = -0.259$, SE $= 0.071$, $P < 0.001$). Under complex task conditions, the conditional indirect effect was positively significant ($\omega_{M1} = 0.371$, 95% CI, $(0.158, 0.624)$), but not significant under simple task conditions ($\omega_{M1} = 0.018$, 95% CI, $(-0.043, 0.090)$); In addition, the index of moderated mediation was significant and positive ($\omega_{M1} = -0.319$, 95% CI, $(-0.551, -0.118)$). Inspirational stimulation showed a different pattern. We found that collaborating with constrained-responsive LLM (compared to collaborating with batch-responsive LLM) negatively influenced cognitive stimulation in simple tasks ($\beta = -0.907$, SE $= 0.182$, $P < 0.001$). Inspiration stimulation was positively related to creativity ($\beta = 0.256$, SE $= 0.068$, $P < 0.001$). The indirect effects were negatively significant in simple task conditions ($\omega_{M2} = -0.233$, 95% CI, $(-0.458, -0.084)$), but not significant under complex task conditions ($\omega_{M2} = 0.012$, 95% CI, $(-0.067, 0.103)$), while the index of moderated mediation was positively significant ($\omega_{M2} = 0.244$, 95% CI, $(0.075, 0.502)$).

Experiment 2 demonstrates that constraining the output of the LLM partner can effectively reduce the creative fixation induced by the LLM in complex tasks, thereby enhancing creative performance. However, in simple tasks, the constraint may lead to a decrease in the exposure of external ideas from the LLM partner, resulting in reduced inspiration stimulation and consequently lower creativity.

Discussion

We found that participants who interacted with an LLM partner (vs a human partner) showed a significant decrease in their self-reported creativity, even though their actual creativity (objective score) was not lower. Thus, the lower self-reported creativity is more likely due to a subjective bias against AI as a collaborator. When individuals collaborate with human teammates, the creative outputs of both parties are more easily perceived as the result of joint exploration—human limitations (e.g., flaws in ideas) allow participants to more clearly identify their unique contributions to the collaboration, thereby maintaining confidence in their own creativity. In contrast, when collaborating with AI, the content generated by AI often exhibits a superficial perfection (e.g., logical fluency, high information density) and richness (producing a large number of ideas in a short time). Such traits
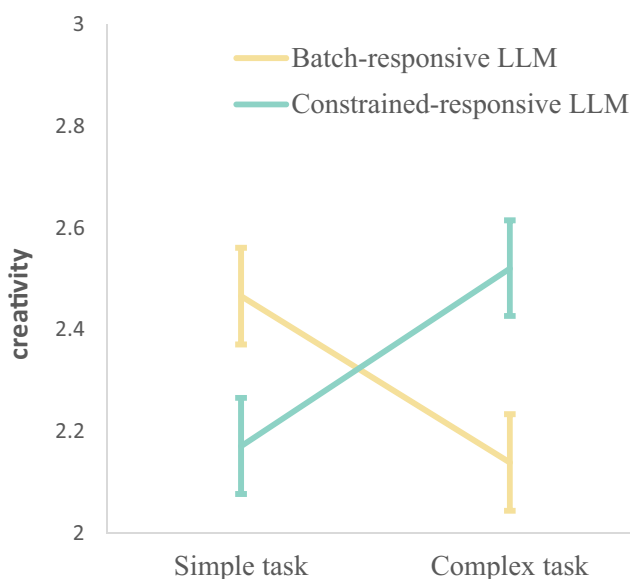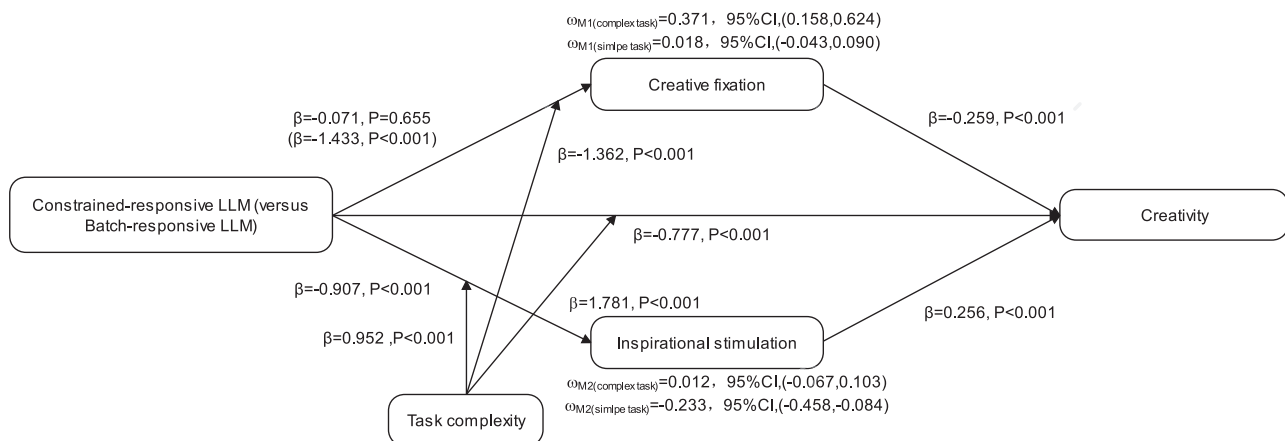

Fig. 5 Moderated mediation. A moderated parallel mediation model demonstrates that constraining the output of the LLM partner can significantly reduce potential creative fixation, particularly in the context of complex tasks. Numbers in parentheses are effects under a complex task. Standard errors are bootstrapped with 10,000 replications.

may implicitly establish a "high benchmark," leading participants to experience a sense of capability gap when comparing their own ideas. Even though the objective quality of participants' ideas does not diminish, the apparent "superhuman performance" of AI still undermines their self-efficacy, making them more conservative in evaluating their own ideas (Bandura 1997).

This work innovatively introduces cognitive load theory to unravel the dual pathways through which LLMs influence creativity and identifies a boundary condition for these pathways. By analyzing objective creativity scores, we found that on the one hand, LLMs enhanced inspirational stimulation, leading to an increase in creativity. On the other hand, they reinforced creative fixation, resulting in decreased creativity. It is precisely due to the existence of such opposing mechanisms that, the main effect of partner type in Experiment 1 falls into the category of a small effect ($\eta_p^2 = 0.024$), but the effect is statistically significant ($P = 0.028$) (Cohen 1988). In fact, this main effect should be understood in conjunction with the significant interaction effect, which is the boundary role of task complexity we emphasized—specifically, in complex tasks, the creative fixation effect induced by LLMs was greater than in the case of simpler tasks, causing a decrease in creativity. We creatively extend the cognitive load theory to the scenario of idea generation supported by LLMs, providing a reference for future exploration of cognitive load in more complex human-AI collaboration scenarios.

The above research results can be further explained through the exploration-exploitation mechanism of constraints. Creativity is fundamentally a dynamic cycle of constraint exploration and exploitation: the exploration phase relies on open associations, and the exploitation phase relies on focused deepening (Tromp 2024). In complex tasks, the large amount of information generated by unconstrained LLMs tends to lead to over-exploration, whereas constrained LLMs promote individuals to enter the exploitation phase through focusing constraints, reducing ineffective exploration, and thereby alleviating creative fixation. In simple tasks, unconstrained LLMs provide rich anchors for the exploration phase, stimulating inspiration, while the focusing effect of constrained LLMs terminates exploration prematurely, thus weakening inspiration stimulation. This confirms that the effectiveness of constraints depends on their alignment with the cognitive phase (exploration or exploitation) required by the task (Tromp 2023).

More importantly, our research contributes to the empirical exploration of the potential negative impact of LLMs on creativity at the individual level. Previous studies have suggested that LLMs may lead to homogenization of ideas at the group level (Anderson et al. 2024; Doshi and Hauser 2024). However, our study observed homogenization at the individual level, namely, creative fixation. This different finding can be attributed to our manipulation of LLM response types and task complexity. Our results indicate that individual creative fixation increases only in the context of collaborating with general, unconstrained LLMs on complex tasks. We believe that the detrimental effects of such fixation accumulate over repeated interactions and, over time, may impair an individual's long-term creative thinking abilities. Secondly, there was no significant difference in the creativity of ideas between the human idea pool and the LLM idea pool in the experiment, and both were at a moderate level. That is to say, we have ruled out the false increase in creativity caused by humans plagiarizing ideas with high creativity scores from their partners. These false effects may mask the negative effects of creative fixation.

In addition, this article offers valuable insights for balancing the utilization of LLMs to enhance creative output while nurturing long-term creativity in individuals. We found that constraining the output of the LLM partner significantly reduced

potential creative fixation, particularly in the context of complex tasks. Conversely, in simple tasks, such constraints resulted in a decline in inspirational stimulation. Based on the above research results, we can tailor suitable LLM response patterns to match different tasks. Specifically, for simple tasks, batch-responsive LLMs can maximize the stimulation of distant associations and inspiration, thereby contributing to the augmentation of creativity. In contrast, for complex tasks, constrained-responsive LLMs allocate more cognitive resources for free thinking, thereby preventing cognitive overload and fixation. These findings provide guidance for the differentiated application of LLMs in a variety of real-world creative tasks.

However, this study has certain limitations. The experiment solely employed ERNIE Bot as the representative LLM. Although ERNIE Bot is well-suited to the Chinese language context and the participants' linguistic background, different models may vary in output characteristics such as diversity and regularity due to differences in training datasets, architectural designs, and training objectives. These variations may affect the performance intensity of inspiration stimulation and creative fixation. For instance, models with strong output diversity may induce less fixation in complex tasks, so the generalizability of the research conclusions needs to be verified with other models. Future research can select different types of models (e.g., general models vs domain-specific models, models with high output randomness vs those with low output randomness) to conduct comparative experiments, systematically examine how model-specific factors regulate the dual mechanisms, and explore the applicable conditions of constrained-responsive interventions in different models. For example, adjusting the constraint intensity for models with flexible output to avoid weakening the effect of inspiration stimulation, and optimizing constraint methods for models that are prone to inducing fixation. These efforts will further clarify the generalizability of the proposed framework for the differentiated application of LLMs and provide more precise references for the selection and optimization of LLMs in different scenarios.

This article also raises other questions worth further exploration. Firstly, we suggest exploring other novel boundary conditions besides task complexity. Other characteristics, such as task format (texts or images) and industry-specific attributes (marketing plan generation, product design, literary composition, etc.), may have different impacts. In addition, our exploration of human-LLM collaboration patterns in creative tasks is preliminary. Future research could explore adjusting the timing of LLMs' outputs (allowing human partners to decide when LLMs' outputs are generated) and the style of collaboration (whether LLMs provide direct answers or guidance-based responses) to match various task characteristics and achieve optimal collaborative creativity. At the same time, we call for the personalized deployment of LLM-based creativity support tools, tailored to the creative task characteristics and tool requirements of diverse industries. This approach is crucial as different task characteristics may activate distinct influence mechanisms.

From the perspective of constraint, future research can advance in three directions. First, optimize the dynamic design of LLM constraints by matching different types of constraints to different stages of tasks—for example, using exclusionary constraints to break thinking stereotypes in the early stage of complex tasks and switching to focusing constraints to deepen solutions in the later stage (Tromp 2024). Second, based on the concept of constraint leveraging power, develop tools to assess individuals' ability to utilize LLM constraints, verify whether it moderates the effects of constraints, and provide a basis for personalized AI assistance (Tromp 2023). Third, move beyond the current single form of constraint based on quantity limits, and test the three-dimensional effect of "constraint intensity × task complexity ×

individual leveraging power" to identify optimal constraint configurations in different contexts, promoting the precise application of LLMs in creative tasks (Tromp and Sternberg 2022).

In summary, our findings indicate that LLMs can improve people's creative performance in simple tasks, but this positive effect is reversed in complex tasks. We emphasize how LLMs influence human creativity through dual, opposing processes (inspiration, stimulation, and creative fixation). The ultimate impact of LLMs on human creativity depends on the relative magnitude of these two effects. In addition, we found that constraining the output of LLMs can mitigate the creative fixation they induce in complex tasks, which has made preliminary efforts to optimize the human-LLM collaboration patterns in creative tasks.

Data availability

References

Anderson BR, Shah JH, Kreminski M (2024) Homogenization effects of large language models on human creative ideation. In: Proceedings of the 16th conference on creativity and cognition. Association for Computing Machinery, pp 413–425

Bandura A (1997) Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 84:191–215

Beguš N (2024) Experimental narratives: a comparison of human crowdsourced storytelling and AI storytelling. Humanit Soc Sci Commun 11:1392. https://doi.org/10.1057/s41599-024-03868-8

Böttger T, Rudolph T, Evanschitzky H, Pfrang T (2017) Customer inspiration: conceptualization, scale development, and validation. J Mark 81:116–131

Bouschery SG, Blazevic V, Piller FT (2024) Artificial intelligence-augmented brainstorming: how humans and AI beat humans alone. Available via https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4724068

Boussioux L, Lane JN, Zhang M, Jacimovic V, Lakhani KR (2024) The crowdless future? Generative AI and creative problem solving. Organ Sci 35:1589–1607

Cillo P, Rubera G (2024) Generative AI in innovation and marketing processes: a roadmap of research opportunities. J Acad Mark Sci 0:1–18

Cohen J (1988) Statistical power analysis for the behavioral sciences, 2nd edn. Routledge, New York. https://doi.org/10.4324/9780203771587

Conroy G (2024) Do AI models produce more original ideas than researchers? Nature. https://doi.org/10.1038/d41586-024-03070-5

Crilly N (2019) Creativity and fixation in the real world: a literature review of case study research. Des Stud 64:154–168

Dennis AR, Valacich JS (1993) Computer brainstorms: more heads are better than one. J Appl Psychol 78:531

DeSanctis G, Gallupe RB (1987) A foundation for the study of group decision support systems. Manag Sci 33:589–609

Doshi AR, Hauser OP (2024) Generative AI enhances individual creativity but reduces the collective diversity of novel content. Sci Adv 10:eadn5290

Eapen T, Finkenstadt DJ, Folk J, Venkataswamy L (2023) How generative AI can augment human creativity. Harv Bus Rev 101:55–64

Elen J, Clark RE (2006) Handling complexity in learning environments: theory and research. Emerald Group Publishing Press

Epstein Z et al. (2023) Art and the science of generative AI. Science 380:1110–1111

Extance A (2023) ChatGPT has entered the classroom: how LLMs could transform education. Nature 623:474–477

Gallupe RB et al. (1992) Electronic brainstorming and group size. Acad Manag J 35:350–369

Hayes AF (2015) An index and test of linear moderated mediation. Multivar Behav Res 50:1–22

Hofstetter R, Dahl DW, Aryobsei S, Herrmann A (2021) Constraining ideas: how seeing ideas of others harms creativity in open innovation. J Mark Res 58:95–114

Hwang AHC, Won AS (2021) IdeaBot: investigating social facilitation in human-machine team creativity. In Proceedings of the CHI conference on human factors in computing systems. Association for Computing Machinery, New York, pp 1–16

Janger J, Schubert T, Andries P, Rammer C, Hoskens M (2017) The EU 2020 innovation indicator: A step forward in measuring innovation outputs and outcomes? Res Policy 46:30–42

Janssen J, Kirschner PA (2020) Applying collaborative cognitive load theory to computer-supported collaborative learning: towards a research agenda. Educ Tech Res Dev 68:783–805

Jia N, Luo X, Fang Z, Liao C (2024) When and how artificial intelligence augments employee creativity. Acad Manag J 67:5–32

Karau SJ, Williams KD (1993) Social loafing: a meta-analytic review and theoretical integration. J Pers Soc Psychol 65:681

Kirschner F, Paas F, Kirschner PA (2011) Task complexity as a driver for collaborative learning efficiency: the collective working-memory effect. Appl Cogn Psychol 25:615–624

Kirschner PA, Sweller J, Kirschner F, Zambrano RJ (2018) From cognitive load theory to collaborative cognitive load theory. Int J Comp Support Collab Learn 13:213–233

Lee BC, Chung J (2024) An empirical investigation of the impact of ChatGPT on creativity. Nat Hum Behav 8:1906–1914. https://doi.org/10.1038/s41562-024-01953-1

Liu Q, Zhou Y, Huang J, Li G (2024) When ChatGPT is gone: creativity reverts and homogeneity persists. Preprint at https://doi.org/10.48550/arXiv.2401.06816

Lu JG, Akinola M, Mason MF (2017) "Switching On" creativity: task switching can increase creativity by reducing cognitive fixation. Organ Behav Hum Decis Process 139:63–75

McGill TJ, Klobas JE (2009) A task-technology fit view of learning management system impact. Comput Educ 52:496–508

Nakadai R, Nakawake Y, Shibasaki S (2023) AI language tools risk scientific diversity and innovation. Nat Hum Behav 7:1804–1805

Osborn AF (1953) Applied imagination. Charles Scribner Press, New York

Pinsonneault A, Barki H, Gallupe RB, Hoppen N (1999) Electronic brainstorming: the illusion of productivity. Inf Syst Res 10:110–133

Potter RE, Balthazard P (2004) The role of individual memory and attention processes during electronic brainstorming. MIS Q 28:621–643

Preacher KJ, Hayes AF (2008) Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. Behav Res Methods 40:879–891

Preacher KJ, Rucker DD, Hayes AF (2007) Addressing moderated mediation hypotheses: theory, methods, and prescriptions. Multivar Behav Res 42:185–227

Rafner J, Beaty RE, Kaufman JC et al. (2023) Creativity in the age of generative AI. Nat Hum Behav 7:1836–1838

Salimzadeh S, He G, Gadiraju U (2024) Dealing with uncertainty: understanding the impact of prognostic versus diagnostic tasks on trust and reliance in human-AI decision making. In: Proceedings of the CHI conference on human factors in computing systems. Association for Computing Machinery, pp 1–17

Shin H, Cotter KN, Christensen AP, Silvia PJ (2020) Creative fixation is no laughing matter: the effects of funny and unfunny examples on humor production. J Creat Behav 54:487–494

Stephen AT, Zubcsek PP, Goldenberg J (2016) Lower connectivity is better: the effects of network structure on redundancy of ideas and customer innovativeness in interdependent ideation tasks. J Mark Res 53:263–279

Sweller J (2010) Element interactivity and intrinsic, extraneous, and germane cognitive load. Educ Psychol Rev 22:123–138

Tromp C (2023) Integrated constraints in creativity: foundations for a unifying model. Rev Gen Psychol 27:41–61

Tromp C (2024) Creativity from constraint exploration and exploitation. Psychol Rep 127:1818–1843

Tromp C, Sternberg RJ (2022) How constraints impact creativity: an interaction paradigm. Psychol Aesthet Creat Arts 18:948–959

Urban M et al. (2024) ChatGPT improves creative problem-solving performance in university students: an experimental study. Comput Educ 215:105031

Acknowledgements

Author contributions

Conceptualization: XC and LZ. Data curation: XC and LZ. Formal analysis: XC and LZ. Investigation: XC and LZ. Methodology: XC and LZ. Project administration: XC and LZ. Visualization: XC and LZ. Writing—original draft: XC and LZ. Writing—review and editing: XC and LZ.

Competing interests

The authors declare no competing interests.

Ethical approval

Approval was obtained from the ethics committee of Renmin University of China on 01 March 2025. The procedures used in this study adhere to the tenets of the Declaration of Helsinki.

Informed consent

This study obtained informed consent from all participants. Informed consent was obtained in writing through the Wenjuan Star platform before the experiment. The consent process was initiated on 29/05/2024, and participants were required to review the details of the study and provide their agreement to participate by selecting the consent option provided within the online questionnaire. Participants in this research have been fully informed of the nature and purpose of the study. All data collected will be used solely for academic purposes, and participants' anonymity is fully assured. Their responses will remain confidential, with no personally identifiable information being collected or stored. There are no anticipated risks associated with their participation, and they were informed of their right to withdraw from the study at any time without any consequences.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1057/s41599-025-05867-9.

Correspondence and requests for materials should be addressed to Xusen Cheng or Lulu Zhang.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.